KU LEUVEN

FACULTY OF ECONOMICS AND BUSINESS

FACULTY OF ECONOMICS AND BUSINESS

THE USE AND CONSEQUENCES OF SUBJECTIVE PERFORMANCE EVALUATION IN SUPERVISOR-EMPLOYEE RELATIONSHIPS

Dissertation presented to obtain the degree of
Doctor in Business Economics
by

Tim Hermans

Number 600 2018

Daar de proefschriften in de reeks van de Faculteit Economie en Bedrijfswetenschappen het persoonlijk werk zijn van hun auteurs, zijn alleen deze laatsten daarvoor verantwoordelijk.

Doctoral Committee

Supervisors

Prof. dr. Alexandra Van den Abbeele	KU Leuven	supervisor
Prof. dr. Martine Cools	KU Leuven	co-supervisor

Members

Prof. dr. Jasmijn Bol	Tulane University
Prof. dr. Eddy Cardinaels	KU Leuven
Prof. dr. Sophie De Winne	KU Leuven
Prof. dr. Victor Maas	University of Amsterdam
Prof. dr. Eelke Wiersma	Vrije Universiteit Amsterdam

Chairman

Prof. dr. Nico Dewaelheyns	KU Leuven

Acknowledgments

Writing this doctoral dissertation has been an interesting and challenging journey. Luckily, it was not a trip in solitude. The completion of this project is due to a large number of companions, who all contributed in their own way to its success. Without their enthusiasm, energy, encouragement, help and support I could not have finished this work. Therefore, they all deserve my utmost gratitude.

Firstly, I want to thank my supervisors, Alexandra and Martine. After a brief taste of the private job market, I realized I wanted to pursue an academic career. I am very grateful to Alexandra and Martine that they provided me with the opportunity to start a PhD. in management accounting. They stimulated me to participate in courses, office meetings, doctoral colloquia and conferences with internationally renowned researchers. Their enthusiasm and continuous support encouraged me to continue with the PhD. and to develop the skills that allowed me to proceed with an academic career. Their guidance and feedback largely improved this dissertation. Thank you for sharing your knowledge and experience with me.

Secondly, I would also like to express my gratitude to my committee members: Jasmijn, Eddy, Sophie, Victor and Eelke. I am very glad you agreed to join my doctoral committee. You are all very busy people and yet you put a lot of time and effort into my dissertation. You provided me with constructive criticism, invaluable suggestions and helpful and thoughtful feedback. Your comments and questions challenged me to reevaluate my research and they undoubtedly contributed substantially to the quality of this work. Thank you for your refreshing insights.

Thirdly, I would like to thank all my colleagues, former and current, at the accounting, finance and insurance department in Leuven and all colleagues at the Antwerp campus. Thank you for the positive atmosphere at work, during seminars, coffee breaks, lunch breaks or after-work activities (Vincent and Alona, thank you for organizing such a fun activities). I could always count on you all for distraction from sometimes frustrating work days and I really enjoyed changing thoughts with you. You contributed a lot to the joy I experienced every day when coming to work.

Fourthly, some people deserve a special mention. Karen, Katlijn and Sharon, my experimental management accounting colleagues, we attended a lot of doctoral courses, Limperg courses or conferences together and I really enjoyed your company there. I enjoyed the experiment and research related talks about Z-tree programming or SPSS issues as well as talks related to more silly topics. That really helped me forward through the doctorate. Alona, Kathleen, Lisa, Nathalie, Roel, Vincent and Yannick, my desk mates at the Antwerp campus, I am glad I shared a desk with you all. You, pleasant group of people, were the major reason I looked forward to come to work every day. I had some great fun with our conversations, both work-related and not work-related. Our interactions helped me to put problems and frustrations into perspective and guided me throughout the PhD. Raf, beacon of wisdom, your door was always open to me. You calmly listened to my questions, frustrations or complaints and you

usually provided my with a clever answer or word of advice. I consider you a very wise and friendly colleague and a great friend!

Fifthly, I would like to thank my parents. They supported me in every single step I took in my life. Their love and support encouraged me to study at KU Leuven, which ultimately led to this doctorate.

Lastly, last but not least, I would like to thank my family. Despite my love for the PhD and the pleasance at work, I always enjoyed coming home to my family after a day at the office. I cannot thank my wife and soulmate Leen enough for her support and presence throughout my life in general and the PhD process more specifically. She is the only one who truly knows me, at my highs and my lows, and she is the one who certainly suffered more as I did during the PhD process. Thank you for your optimism, encouragement, support and patience with me. You bring both peace and pleasure in my life. Furthermore, you gave birth to my son Jonathan, whose cheerful and playful presence brightens each and every day. He constantly reminds me of what is really important in life. Finally, our unborn child, I cannot wait to meet and welcome you into the family!

Tim Hermans
Kontich, January 2018

Comments and Conferences

The first chapter of this PhD dissertation consists of a literature review of the academic research on subjective performance measurement. The literature review discusses 67 articles in 20 high-impact journals over the period 1977 to 2013 and has been published in the Review of Business and Economic Literature (Hermans et al., 2013). The second, third and fourth chapter consist of research papers co-authored with Martine Cools and Alexandra Van den Abbeele. All three research papers in this PhD dissertation greatly improved by the feedback of my supervisors and committee members: Jasmijn Bol, Eddy Cardinaels, Martine Cools, Sophie De Winne, Victor Maas, Alexandra Van den Abbeele and Eelke Wiersma, as well as many other people.

For the second chapter, we received helpful suggestions from Ariela Caglio, Stephan Kramer, Evelien Reusen, Karl Schuhmacher, Michael Shields, Kristy Towry, Wim Van der Stede, Martijn van der Steen, Sally Widener, and participants at the 2014 Limperg course in Experimental Research in Accounting at Tilburg University, at the European Network for Research on Organizational and Accounting Change 7th Doctoral Summer School in Management Accounting, at the 2015 EAA European Doctoral Colloquium in Accounting, at the 2016 conference on New Directions in Management Accounting: Innovations in Practice and Research, at the 2017 Research Day in Accounting and at the 2017 European Network for Experimental Accounting Research conference.

For the third chapter, we received helpful suggestions from John Christensen, Thomas De Groot, Henri Dekker, Kathryn Kadous, Karl Schuhmacher, Marcel Van Rinsum, and participants at the 2015 European Network for Experimental Accounting Research Summer School, at the Amsterdam Research Center in Accounting seminars 7th of March 2016 at Vrije Universiteit Amsterdam, at the 2016 Annual Conference for Management Accounting Research and at the 2016 Annual Conference of the European Accounting Association.

Table of Contents

Table of Contents

List of Figures

List of Tables

General Introduction

In this PhD dissertation, I investigate the use and consequences of subjective performance evaluation in supervisor-employee relationships from a management control perspective. The first section of this introduction describes the general research motivation. The second section discusses the research method used in this dissertation. The third section provides an overview of the different chapters in this dissertation and their interrelatedness.

Research Motivation

Management control systems (MCS) include all systems internal decision makers in organizations use to help ensure that organizational strategies and goals are implemented (Horngren et al., 2015; Merchant & Van der Stede, 2007; Salterio, 2015). Amongst others, MCS measure, analyze and report financial as well as nonfinancial information that is meant to be useful to managers in order to perform their job well (Horngren et al., 2015; Otley. 1999; Sprinkle, 2003; Sprinkle & Williamson, 2007). In this dissertation, we focus on a specific type of management control, results control, and more particularly on subjective performance evaluation.

Results control concerns the measurement and evaluation of performance aimed to motivate employees to generate the outcomes the organization wants (Merchant & Van der Stede, 2007). Performance evaluation and rewarding are essential control instruments for organizations to increase employee productivity and goal congruence between employee and organization (Baiman & Rajan, 1995; Ferreira & Otley, 2009; Höppe & Moers, 2011). More and more firms introduce employee incentive programs (WorldatWork & Vivient Consulting, 2012), but it is hard to capture employees' individual contribution to firm value based on only objective performance measures generated by the accounting system (Bol, 2008; Maas et al., 2012). Subjective performance evaluation is a common practice in organizations in which a manager receives the decision-making power to evaluate and reward the performance of an employee subjectively (Choi et al., 2016; Murphy & Oyer, 2003; Rajan & Reichelstein, 2006). Managers can then complement objective measures with subjective assessments of performance (Baker et al., 1994) or they can determine subjectively the specific weightings placed on the various performance dimensions (Ferreira & Otley, 2009). The correctness of a subjective evaluation cannot be verified by a third party, because it entails subjective judgment, personal impressions or opinions or additional, private performance information that is not easily contractible, controllable or foreseeable (Ahn et al., 2010; Bailey et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; Choi et al., 2016; Gibbs et al., 2004; Rajan & Reichelstein, 2009). Next to the benefits, there are some disadvantages attached to subjectivity in performance contracts as well (Bol, 2008). Judgment biases – or perceptions of them – , favoritism and evaluation uncertainty may harm employee motivation and productivity and may cause conflicts between managers and employees (Ahn et al., 2010; Bol, 2011; Ferreira & Otley, 2009; Golman & Bhatia, 2012; Ittner et al., 2003; Lipe & Slaterio, 2000; Moers, 2005;

Prendergast & Topel, 1993). Furthermore, subjective performance evaluations can be time-consuming and manager's time is scarce and costly (Bol, 2011; Ferreira & Otley, 2009).[1]

This doctoral research project aims to provide further insights into how managers make use of their decision-making power to evaluate and reward the performance of an employee subjectively. We start this doctoral dissertation with an extensive overview of academic research on subjective performance evaluation (chapter 1). Next, we research how managers use their decision-making power depending on MCS design elements such as: manager-employee compensation inequality and managerial discretion to freely decide on the bonus size (chapter 2); the accuracy of the performance information on which managers base their evaluations and whether (or not) managers get the opportunity to write a justification on how they allocated employee bonuses (chapter 3), and relative performance information that measures performance of an employee in relation to that of one or more other employees (chapter 4). We further aim to advance knowledge on the consequences of managers' use of their decision-making power for employee performance, employees' trust in the manager, managers' fairness perceptions and managers' bonus allocations as well.

This research is highly relevant for practice, as a recent survey indicates that two-thirds of the private companies use subjectivity in their bonus plans (WorldatWork & Vivient Consulting, 2012). The dissertation is relevant for academia as well. Salterio reviewed and compared management accounting research over the periods 1990-1993 and 2010-2013 in six major accounting journals (Salterio, 2015). He concluded that *"Areas of emerging knowledge interest in management accounting include performance measures, performance evaluation systems,[…]. Areas of ongoing interest, which suggest that the knowledge base has not yet satisfied demand for knowledge, include effects of incentives (both empirically and via modeling in principal-agent models) and control systems (both in the effects on decision makers and the development of such systems)."* (Salterio, 2015). This dissertation aims to contribute to these areas of management accounting literature. Further, scholars indicate the importance of empirical studies on how (elements of) the MCS and information produced by the MCS affect the behavior and decisions of individuals (Sprinkle, 2003) and how the MCS might help to improve judgments and decisions (Sprinkle & Williamson, 2007). The effect of MCS elements on judgments, decisions and performance can be moderated by various individual, task and environmental variables (Bonner & Sprinkle, 2002; Luft & Shields, 2003).

Theoretical agency models on performance evaluations assume decision-makers are perfectly rational creatures seeking for optimal decisions serving their self-interest (Baiman & Rajan, 1995; Baker et al., 1994; Budde, 2007; Golman & Bhatia, 2012; Levin, 2003; MacLeod, 2003; Prendergast & Topel, 1996; Rajan & Reichelstein, 2006, 2009), but there is a lot of evidence that indicates decisions made by actual, 'real' users of information deviate from rationality (Bonner, 1999; Krishnan et al., 2005; Sprinkle, 2003; Sprinkle & Williamson, 2007). Subjective performance evaluations are subject to plenty of judgement biases such as common measure bias, compression bias, gender and race biases, leniency bias,…(Ahn et al., 2010; Biernat & Sesko, 2013; Bol, 2011; Bol et al., 2016; Elvira & Town, 2001; Golman &

[1] We refer to chapter 1 of this dissertation for a more extensive review of the academic research related to subjective performance evaluation and its benefits and costs.

Bhatia, 2012; Lipe & Salterio, 2000, 2002; Moers, 2005; Prendergast, 1993) and simple heuristics such as anchoring (Bailey et al., 2011). It is vital for organizations to understand how managers weigh and combine various performance measures into an overall subjective performance evaluation, but it is unclear how this process works and which elements influence the weights placed on the various performance measures (Sprinkle, 2003; Sprinkle & Williamson, 2007). Understanding these elements and processes is important in order to improve judgment and decision-making (Bonner, 1999).

Decision-makers follow, next to economic incentives, ethical, social and moral principles such as honesty, fairness, equity or reciprocity (Adams, 1963, 1965; Bolton & Ockenfels, 2000; Cox, 2004; Falk & Fischbacher, 2006; Fehr et al., 1997; Fehr & Gächter, 2000; Fehr & Schmidt, 1999, 2004; Hargreaves Heap et al., 2013; Itoh, 2004; Luft, 1997; Maas et al., 2012; Malhotra & Murnighan, 2002; Rabin, 1993; Smith, 2011). Other-regarding preferences or reputational considerations could increase or reduce the need for certain management accounting practices and thus they may have implications for MCS design and practices and the use of MCS information (Sprinkle, 2003; Sprinkle & Williamson, 2007). Previous research indicates that formal control systems can enhance cooperation and reciprocity (Coletti et al., 2005), but they can reduce reciprocal co-operation as well (Tenbrunsel & Messick,1999). Additional research is needed on when formal control systems amplify or demolish reciprocity and cooperation and which elements of the MCS might improve judgments and decisions, might increase or decrease the occurrence of judgment biases and might elicit other-regarding preferences (Sprinkle & Williamson, 2007). One of the main goals of MCS is to temper the conflict of interest between employees and managers and to motivate employees to maximize firm value. Current research overly focuses on employee moral hazard and neglects managerial moral hazard (Sprinkle, 2003; Sprinkle & Williamson, 2007). However, due to subjective performance evaluations, employees are very vulnerable to managers that misinterpret performance measures and that allocate bonuses to their own advantage (Fisher et al., 2005). An important avenue for future research lies in understanding whether or when individuals act opportunistically and how other-regarding (social, ethical,…) motives interact with formal MCS in order to help mitigate conflicts between managers and employees (Fisher et al., 2005; Sprinkle & Williamson, 2007).[2]

Finally, experimental research in management accounting typically focuses on simple one-period settings, with single, one-dimensional tasks performed by a single person (Sprinkle & Williamson, 2007). Additional research is wanted given the more complex nature of many decision-making situations in practice. Research could focus on the dynamic (multi-period) effects of MCS and manager's subjective performance evaluation on employee motivation and effort. Reputational considerations in multi-period settings may substitute or complement formal management accounting practices (Kreps et al., 1982; Sprinkle & Williamson, 2007). Additionally, research could focus on how MCS motivate their employees to allocate high levels of effort to multiple tasks, responsibilities or performance dimensions without inducing disproportionate effort allocations to certain tasks, responsibilities or performance dimensions

[2] We refer to Luft (1997) for a thorough discussion regarding how other-regarding preferences such as fairness and ethical concerns affect management accounting practices (Sprinkle & Williamson, 2007).

(Brüggen & Moers, 2007; Hannan et al., 2013; Hecht et al., 2012; Sprinkle, 2003; Sprinkle & Williamson, 2007). Lastly, research could look at how MCS motivate cooperation and competition among different employees (Hannan et al., 2013; Luft, 2016; Sprinkle, 2003). In our three research studies we aim to contribute to the research issues discussed above.

Research Method: Laboratory Experiments

In all three research papers we conducted laboratory experiments in order to answer our research questions. *"An experiment is a scientific investigation in which (independent) variables are manipulated and their effects on (dependent) variables are observed."* (Sprinkle, 2003; Sprinkle & Williamson, 2007). Experiments are frequently used in management accounting research. Interested readers can consult Luft & Shields (2003) for an extensive overview of this extant literature (Sprinkle & Williamson, 2007). Experiments allow to create a research setting and generate data fit for the particular research question. They allow for a ceteris paribus change in particular aspects of the performance evaluation system which allows to isolate specific causes in order to investigate their theoretical effects (Sprinkle, 2003; Sprinkle & Williamson, 2007). Researchers can control for alternative explanations present in the natural environment through the manipulation of the independent variables, through keeping constant some elements in the research setting, through the randomization of participants over experimental treatments or through the collection of a rich set of alternative explaining variables measured as required by the experiment (Bonner, 1999; Sprinkle & Williamson, 2007). This research method is especially capable of measuring and documenting the process through which decisions are made (Sprinkle & Williamson, 2007). This is essential for organizations that want to understand how managers weigh and combine various performance measures into an overall subjective performance evaluation. So far it is unclear how this process works and which elements influence the weights placed on the various performance measures (Sprinkle & Williamson, 2007).

Well-conducted experiments allow for a thorough test of theoretical predictions, because they have strong internal validity. Therefore, experiments are particularly useful to investigate cause-effect relations between independent and dependent variables (Sprinkle & Williamson, 2007). Experiments even allow to conduct research on issues that do not exist yet in the "real world" (Bonner, 1999). However, this is experiments' weakness at the same time. Experiments often lack external validity. The stylized, simplified research settings do not take into account all relevant aspects of reality and by doing so they sacrifice some external validity. The representativeness and generalizability of experiments is often questioned (Bonner, 1999; Sprinkle & Williamson, 2007).

Archival or survey data contains more mundane realism. However, archival or survey data on management accounting issues such as a manager's individual decision-making process is often difficult or impossible to obtain (Bonner, 1999; Sprinkle & Williamson, 2007). Furthermore, participants featuring in these data sets are not random; these data sources suffer from self-selection or sample-selection biases. Next, dependent and independent variables under investigation may be measured imprecisely, which can generate measurement error. Further, the results may be confound by other alternative explanations that cannot be unraveled, because there was no data collected on those alternative explanations (omitted variables bias)

and the researcher was not able to control or observe the environment. As such, archival or survey data allows only to make correlational interferences about the relationship between independent and dependent variables instead of causal interferences (Bonner, 1999). All these issues with archival or survey data can endanger the internal validity, construct validity and statistical conclusion validity of the research (Sprinkle, 2003; Sprinkle & Williamson, 2007). The strength of experiments lies in their internal validity, their capability to make causal interferences and to rigorously test theoretical predictions and to refine theory. Experiments are well-suited to provide evidence regarding why actual behavior differs from behavior predicted by economic agency models (Sprinkle, 2003).

Overview of the Four Chapters

To advance the literature on subjective performance measurement, this doctoral project consists of a literature review on this topic (chapter 1) as well as three experimental papers (chapters 2-4). These empirical papers provide better insights and concrete recommendations to organizations on how to design their performance evaluation and rewarding system in order to stimulate employee productivity and goal congruence.

Chapter 1

The first chapter of this dissertation consists of an extensive literature review we undertook during the first year of the PhD. project. The article was published in the Review of Business and Economics (Hermans et al., 2013). It provides an overview of the academic literature on subjective performance measurement in high-impact journals during the period 1977–2013 and identifies a number of research opportunities, which we address in the papers of the doctoral research project.

Chapter 2

The second chapter investigates, within manager-employee dyads, how managerial discretion to freely decide on the bonus size of an employee (high, low) and manager-employee compensation inequality (high, low) affect employee effort, managers' bonus allocation and the extent to which a manager is concerned about a fair bonus allocation. Using a 2x2 experiment, we develop and find support for a causal model that explains how managers subjectively allocate a bonus within this context. More specifically, management control systems that limit managerial discretion prevent that managers act overly opportunistic and therefore cause higher bonuses for the employee compared to control systems that do not limit managerial discretion. Furthermore, control systems that limit managerial discretion improve employee effort, and managers reciprocate to this extra employee effort by offering a higher bonus to the employee. However, in case of manager-employee compensation inequality the presence of a control system that limits managerial discretion provides a manager with a legitimate excuse to act within the boundaries of the control system. As a result, focusing on long-term self-interest becomes justifiable, in contrast to paying more attention to fairness. This increased focus on long-term self-interest instead of fairness will ultimately lead to lower bonus allocations to the employee.

Chapter 3

The third chapter examines, in a multi-employee setting, how the accuracy of the performance information on which managers base their evaluations (high, low) and whether managers get the opportunity to write a justification on how they allocated employee bonuses (present, absent) affect managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations and managers' compression in bonus allocations. Using a 2x2 experiment, we elaborate and provide evidence for a causal model that explains how managers subjectively allocate a bonus within this context. Higher (perceived) information accuracy increases differentiation in bonus allocations compared to lower (perceived) information accuracy. Furthermore, managers estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations as higher when the performance information accuracy is (perceived as) high and when they are allowed to justify the bonus allocation to the employees compared to situations where either information accuracy or a possibility for justification is missing. Managers will consequently differentiate more in their bonus allocations to employees, in line with their estimates of employees' perceived procedural fairness.

Chapter 4

The fourth chapter examines, in a multi task context, how task complexity (high, low) and employees distortion of effort allocations across tasks away from the firm-preferred equal effort allocation affect employee performance. Additionally, we investigate how a management control system consisting of detailed relative performance information and a discretionary bonus system affects employees' effort allocation across tasks. Using two experiments, we demonstrate that a distorted allocation of effort across tasks reduces employee overall performance. However, the negative effect of the distorted allocation of effort across tasks on overall employee performance is less negative for complex tasks than for simple tasks. Concerning employees' effort allocation across tasks, we find that, in the first period, employees will focus on the task for which they have the highest skills. Employees will focus on the easiest task in order to perform at least well on that task. However, after the provision of relative performance information and the allocation of the discretionary bonus, employees reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague.

Relation between chapters

As indicated earlier, management control systems (MCS) include all systems internal decision makers in organizations use to help ensure that organizational strategies and goals are implemented (Horngren et al., 2015; Merchant & Van der Stede, 2007; Salterio, 2015). Any assessment of the role of MCS information starts from considering how managers make use of the information being provided to them (Otley. 1999). The particular use of the control information can be more important than the formal design of the control system itself (Ferreira & Otley. 2009). In this dissertation, we therefore focus on how decision makers make use of the information in a performance evaluation context and how this information and the MCS influence the behavior of the decision maker.

One of the main goals of MCS is to temper the conflict of interest between employees and managers and to motivate employees to maximize firm value. Chapters 2-4 focus on issues in manager-employee relationships that obstruct a good implementation of organizational strategies and goals and how a deliberate choice for particular elements of the management control system can improve that implementation. Chapter 2 and 3 focus on two important problems with managers' behavior that negatively affect employee and organizational performance: managers that act opportunistically and managers that do not differentiate in their bonus allocations. Chapter 4 focuses on the potential lack of goal congruence between organization and employee which might lead to problematic employee behavior that negatively affects employee and organizational performance. We identify MCS elements that enable managers (chapter 2 and 3) and employees (chapter 4) to make better decisions (Sprinkle, 2003).

Furthermore, it is important for organizations to understand how managers come to their overall subjective performance evaluation in order to improve judgment and decision-making. The evaluation process and the elements that influence managers' decision are still unclear (Bonner, 1999; Sprinkle, 2003; Sprinkle & Williamson, 2007). In chapters 2 and 3 we try to disentangle the managers' evaluation process by means of a path model. We demonstrate for instance that managers follow both economic incentives and ethical, social and moral principles such as fairness, equity or reciprocity in their decision making. Furthermore, the extent to which managers consider ethical, social or moral principles depends on the specific combination of management control elements such as manager-employee compensation inequality and managerial discretion to freely decide on the bonus size (chapter 2) or the accuracy of the performance information on which managers base their evaluations and whether managers get the opportunity to write a justification on how they allocated employee bonuses (chapter 3).

Chapter 1
Subjective Performance Measurement: A Literature Review

Tim Hermans[3], Martine Cools and Alexandra Van den Abbeele

Abstract
This article provides an overview of academic research on subjective performance measurement, a practice that intends to remedy the weaknesses of evaluations solely based on objective quantitative performance measures. The literature on subjective performance measurement mainly focuses on four research streams: optimal contracting, discretionary bonus pools, judgment biases and debiasing, and perceived fairness. We discuss these four research streams as encountered in 67 articles published in 20 high-impact journals over the period 1977 to 2013. In addition, this article identifies several research gaps and avenues for future research.

Keywords: literature review; management control; subjective performance measurement
JEL codes: J33, M52

1.1 Introduction

This article provides an overview of the academic literature on subjective performance measurement. This research field captures the common practice in organizations that a manager or supervisor evaluates the performance of an employee or subordinate subjectively. The subjectivity in performance evaluation can be present in several ways. Supervisors can use subjective performance measures, they can ex post adjust the weighting of objective performance measures and/or they can make discretional adjustments based on factors different from the performance measures specified ex ante (Bol, 2008; Bol & Smith, 2011). The subjective evaluations are based on personal impressions or opinions (Bol & Smith, 2011) or information not explicitly contracted on because it represents unforeseen circumstances that would not be contractible in objective, formula-based performance evaluations (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995). Bommer et al. (1995) indicate that the correlation between objective performance measures and subjective ratings of employee performance is only 0.39. Subjective performance measures are thus clearly distinct from objective performance measures and as such sufficient attention needs to be paid to their design and use. This paper therefore offers an extensive overview of existing research on subjective performance measurement.[4]

[3] Corresponding author, KU Leuven, Campus Antwerp, Korte Nieuwstraat 33, 2000 Antwerp.
E-mail: Tim.Hermans@KULeuven.be, Tel. +32 16 37 62 50.
[4] Our study is not the first one to offer an overview of the subjective performance measurement literature. Bol (2008) examines the role of subjectivity in compensation contracts. She describes optimal contracting in a traditional agency context and thereby depicts the benefits and costs related to subjectivity in compensation

As a research method for this literature review we searched for published articles on subjective performance measurement in the Web of Science. The following search terms were used: 'subjective performance', 'subjective evaluation', 'subjective measurement', 'subjective measure', 'subjective judgment', 'subjective assessment', 'subjective review', 'performance ratings' and 'evaluation'. We investigated whether these search terms occurred either in the topic or in the title of published articles. Publications in journals with an impact factor larger than 1 were retained for further analysis. Afterwards, we screened all obtained articles to make sure they were relevant for the purpose of this literature review.[5] This led to our final sample of 67 articles published in 20 high-impact journals over the period 1977 to 2013. We grouped these 67 articles in 4 research streams already defined in the literature based on the keywords of the most highly cited papers. For instance, Maas et al. (2012) deals with 'optimal contracting', 'discretionary bonus pools' and 'perceived fairness'. Baiman & Rajan (1995) examines 'optimal contracting' and 'discretionary bonus pools'. MacLeod (2003) researches 'optimal contracting' and 'judgment biases and debiasing'. Baker et al. (1994), Ke et al. (1999) and Levin (2003) investigate 'optimal contracting', Gibbs et al. (2004) research 'discretionary bonus pools' and Dulebohn & Ferris (1999) and McFarlin & Sweeney (1992) look into 'perceived fairness'. Libby et al. (2004), Lipe & Salterio (2000, 2002), Moers (2005) and Prendergast & Topel (1993) focus on 'judgment biases and debiasing'. The topics of those highly cited papers resulted in 4 research streams: 'optimal contracting', 'discretionary bonus pools', 'judgment biases and debiasing' and 'perceived fairness'. Afterwards, we were able to fit the remaining papers of our sample in this structure based on their topic or keyword. Some papers address multiple research streams and they therefore reappear in one or more of the subsequent sections discussing each research stream separately. The remaining of this article is organized as follows. In the next sections, we discuss the four broad research streams on subjective performance measurement: optimal contracting (section 2), discretionary bonus pools (section 3), judgment biases and debiasing (section 4) and perceived fairness (section 5). Subsequently, we deal with a number of research opportunities identified through this literature review in section 6 and we end with a conclusion in section 7.

1.2 Optimal Contracting

Traditional academic research in agency theory focuses on objective performance measurement and optimal contracting. In these classical principal-agent models, a principal designs an optimal contract inducing an agent to exert effort that maximizes the value relevant

contracts. Our analysis differs from the analysis of Bol (2008) because we collected a more extensive amount of papers touching more aspects of subjectivity in performance evaluation than contracting alone. Franco-Santos et al. (2012) provide a framework to classify contemporary performance measurement systems and apply this to their review of 76 empirical studies. They discuss perceptions of subjectivity, justice and trust, and judgment biases. In contrast to their general and high-level classification framework for 'all' contemporary performance measurement systems, we provide a more in-depth overview and discussion of the subjective performance measurement literature only. Prendergast & Topel (1993) discuss potential pitfalls of subjective performance evaluations: they review supervisors' preferences and biases such as leniency bias, favoritism and compression bias. We update their observations and extend the scope.

[5] Although our literature review is quite extensive, we focus on subjective performance measurement and therefore do not discuss articles dealing with feedback, performance appraisal, relative performance evaluation, self-evaluation or peer-evaluation.

to the principal. The agent gets rewarded for his effort, but he is effort-averse. The principal cannot fully observe or verify the actions undertaken by the agent and must rely on a number of objective performance measures. Appropriately designed incentive contracts can provide the agent with incentives to act in the interest of the principal and as such optimal incentive contracts mitigate or resolve agency problems (Bol, 2008; Cronqvist & Fahlenbrach, 2013).

With an optimal incentive contract, the principal does not have to monitor the agent's behavior. He can just rely on the objective outcome measures that measure the agent's performance. In other words, an agent's incentive contract provides the principal with a substitute for monitoring the agent's behavior (Morse et al., 2011). At the same time, these incentive contracts transfer risk from the principal to the agent as the objective performance measures used in these contracts do not capture the agent's effort completely and accurately. Indeed, performance in most jobs cannot be measured objectively because joint production makes individual output not readily quantifiable (Baker et al., 1988; Levin, 2003). In addition, the range of possible actions that the agent can take is too extensive to contract upon ex ante (Baker et al., 1988). As such, high-uncertainty environments warrant greater reliance on subjective performance criteria (Keeley, 1977). In practice, objective performance measures are therefore often complemented with subjective performance measures. Table 1.1 provides an overview of published articles on optimal contracting including subjective performance evaluations.[6] The first article by Bol (2008) is a literature review examining the role of subjectivity in compensation contracts. In Table 1.1 we update and extend Bol et al.'s overview. We first discuss the articles that extend the traditional agency theory models to account for subjectivity in performance measurement. Next we include articles revealing the benefits of subjectivity in performance contracts, to end with the articles dealing with the costs related to subjectivity in optimal contracts.

[6] The tables in this article are divided into several topics. The papers in the tables are alphabetically ordered by author name(s) within these topics.

Table 1.1 Optimal contracting with subjective performance evaluations

Article	Focus/Results	Research Set-up
Introduction to optimal contracting		
Bol (2008)	Optimal contracting in traditional agency context: benefits and costs related to subjectivity.	Literature review
Optimal contracting models including subjectivity		
Baiman & Rajan (1995)	Discretionary bonus pools are an efficient way to incorporate non-contractible information in a two-agent setting.	Analytical model
Budde (2007)	Researches a BSC with contractible and not-contractible scorecard measures: a combination of a formal contract and a subjective performance evaluation may outperform a purely formal contract.	Analytical model
Cronqvist & Fahlenbrach (2013)	Private equity sponsors (strong principals) use less subjective performance measures, but some subjective performance evaluation to compensate the CEO.	Field study (CEO contract data on leveraged buyouts of 20 large, American listed firms, 2005-2007)
Höppe & Moers (2011)	Different types of subjectivity are used for different purposes: "discretionary bonuses" are used for risk-reduction, "subjective weights" for congruity-improvement.	Archival study (1,753 firm-year-observations for 424 American, publicly listed firms, 1998-2002)
Ke, Petroni & Safieddine (1999)	Privately held insurers (strong principals) use more subjective performance measures to compensate the CEO.	Archival data (45 privately-held and 18 publicly-held American insurers, 1994-1996)
MacLeod (2003)	Extends standard principal-agent model with a single agent with subjective evaluations.	Analytical model
Rajan & Reichelstein (2006)	When the bonus pool covers many agents and/or the principal's subjective information is precise, discretionary bonus pools are nearly as efficient as explicit contracts.	Analytical model
Rajan & Reichelstein (2009)	In the single-agent case it might be optimal to ignore the subjective signal with discretionary bonus pools.	Analytical model
Benefits of subjectivity in optimal contracts		
Baker, Gibbons & Murphy (1994)	A combination of objective and subjective measures sometime outperforms an explicit or an implicit contract alone.	Analytical model
Baker, Jensen & Murphy (1988)	Discusses several benefits and costs related to objective and subjective performance measurement.	Literature review
Gibbs, Merchant, Van der Stede & Vargus (2004)	Subjective bonuses are used to complement perceived weaknesses in quantitative performance measures and to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998-1999) and 1050 surveys in 326 different dealerships
Höppe & Moers (2011)	"discretionary bonuses" are used for risk-reduction, "subjective weights" for congruity-improvement.	Archival study (1,753 firm-year-observations for 424 American, publicly listed firms, 1998-2002)
Indjejikian & Matejka (2012)	Nonfinancial measures or subjective evaluations are more used for bonuses when the recipients have greater influence over the internal accounting systems design.	Survey (242 BU-managers and controllers of 121 BUs of 7 Dutch multinationals and 48 additional interviews)
Ke, Petroni & Safieddine (1999)	Privately held insurers (strong principals) use more subjective performance measures to compensate the CEO.	Archival data (45 privately-held and 18 publicly-held American insurers, 1994-1996)
Keeley (1977)	High-uncertainty environments warrant greater reliance on subjective performance criteria.	Questionnaire (106 supervisor-subordinate pairs)

Article	Focus/Results	Research Set-up
Costs of subjectivity for optimal contracts		
Ahn, Hwang & Kim (2010)	Subjective measures provide less incentive than objective measures because of the lack of variation in scores (compression bias).	Archival (13 government-invested companies, Republic of Korea, 1990-2006)
Baker, Jensen & Murphy (1988)	Discusses several benefits and costs related to objective and subjective performance measurement.	Literature review
Golman & Bhatia (2012)	Subjective performance evaluation leads to leniency bias, and associated with that reduced employee effort.	Analytical model
Krishnan, Luft & Shields (2005)	Individuals do not weigh measures appropriately in a two-measure incentive system.	Experiment (32 accounting and MBA students)
Levin (2003)	Self-enforced relational contracts with moral hazard result in compression bias.	Analytical model
MacLeod (2003)	Optimal contracts with subjective evaluations can result in compression bias and leniency.	Analytical model
Morse, Nanda & Seru (2011)	Powerful CEOs can shift the weight on performance measures toward the better performing measures. This affects future firm performance negatively.	Archival study with 2348 firms over the period 1992-2003
Prendergast & Topel (1996)	Subjectivity leads to favoritism.	Analytical model
Prendergast (1993)	With subjective evaluation procedures, subordinates have an incentive to conform to what they feel their superiors want to hear.	Analytical model

The first eight articles extend traditional agency theory knowledge with one or several aspects of subjectivity. MacLeod (2003) allows for subjective performance evaluations in the standard principal-agent model by including subjective performance measures. With this analytical model, he shows that if the principal's and the agent's subjective evaluations correspond (or equivalently if there is trust and perceived fairness between principal and agent), one can implement the optimal contract just as if subjective evaluations were objective and verifiable. Budde (2007) provides a theoretical model for a combination of objective and subjective performance measures in a balanced scorecard (BSC) setting. The model shows that when all objective performance measures are perfectly verifiable, a properly designed BSC can perfectly align the interests of the principal and the agents with an explicit contract.

When not all BSC measures are contractible[7], the first-best solution, a contract in which the agent exerts the optimal level of effort that provides the optimal value relevant to the principal, may still be obtained through a combination of a formal contract and a subjective performance evaluation (Budde, 2007). Höppe & Moers (2011) undertook an archival study in which they focus on the use of two different types of subjectivity: "subjective weights" and "discretionary bonuses". "Subjective weights" concern the option whereby supervisors can ex post adjust the weighting of objective performance measures, while "discretionary bonuses" refer to the case where supervisors can make discretional adjustments based on factors different from the performance measures specified ex ante. According to optimal contracting considerations, their

[7] A performance measure is contractible if its value is observable both by the principal, the agent and an unrelated third party. In this case the performance measure can be explicitly incorporated in a contract (Baiman & Rajan, 1995).

results show that “subjective weights” are used to improve goal congruence between the agent and the principal, while “discretionary bonuses” are used to reduce risk for the agent due to uncertainty.

Baiman & Rajan (1995) and Rajan & Reichelstein (2006, 2009) provide analytical models on the use of discretionary bonus pools. For a discretionary bonus pool, the bonus pool is based on an explicit formula involving objective performance measures agreed-upon ex ante. Afterwards, the bonus pool is allocated amongst the agents at the principal’s discretion. The entire bonus pool is paid out regardless of the subjective information observed by the principal, but in case of unfavorable subjective information the principal withholds part of the bonus of one agent to give it to other, better-performing agents. Baiman & Rajan (1995) prove that discretionary bonus pools result in a strict Pareto improvement compared to the optimal contract that does not use non-contractible information by enabling a principal to exploit non-contractible information to motivate agents. Furthermore, Rajan & Reichelstein (2006) show that discretionary bonus pools are optimal when a principal must rely solely on non-verifiable, subjective information to create incentives for a group of agents. They find that bonus pools are nearly as efficient as explicit contracts, provided that the bonus pool covers a large number of agents and/or the principal’s subjective information is fairly precise. In addition, when no other agent is present, the principal incurs an additional cost when the agent shirks. The model of Rajan & Reichelstein (2009) indicates that in the single-agent case it might be optimal to ignore the subjective signal. When both objective and subjective measures are used, the optimal contract results in less divergent performance scores relative to the number of performance levels on the different performance measures than when only objective measures are used. Furthermore, they show that the single-agent bonus pool results in less divergent performance scores relative to the number of possible performance scores on the different performance measures than a multiple-agent bonus pool (Rajan & Reichelstein, 2009).

In agency models extended with subjective performance measures, strong principals have greater incentives to observe and monitor agents’ effort and to base agents’ reward on those subjective observations. Consequently, agent’s compensation is less likely based on an explicit contract with objective performance measures (Ke et al., 1999). Ke et al. (1999) confirm this theoretic reasoning empirically via an archival study amongst privately-held and publicly-held property-liability insurers. They find that within privately-held insurers (called strong principals) CEO compensation is less based on objective measures like accounting information and presumably more on subjective measures compared to the publicly-held insurers (called weak principals). Consequently, their findings are consistent with optimal contracting (Ke et al., 1999). Cronqvist & Fahlenbrach (2013) come to the opposite conclusion in their study of CEO contracts within large American firms moving from public ownership with dispersed owners (weak principal) to private ownership with strong principals. They find that strong principals redesign contracts away from qualitative, nonfinancial measures, but they introduce subjective performance evaluation instead. Baker et al. (1994) assume objective performance measures are imperfect and cause incentive distortions, which can be mitigated by the inclusion of subjective performance assessments. The authors prove that in some circumstances, neither an explicit nor an implicit contract alone yields positive profit, but a combination of objective and subjective measures can.

Next, we discuss seven articles on the benefits of subjectivity in optimal contracts. Gibbs et al. (2004) use archival and survey data on compensation of managers in car dealerships to examine when firms make greater use of subjectivity in bonus payments. It turns out that subjective bonuses are used to respond to perceived weaknesses in quantitative formulaic bonuses such as incompleteness, short-term focus and susceptibility to manipulation. Using – only imperfect – objective performance measures may lead to suboptimal actions taken by the agents (Baker et al., 1994). Agents tend to focus their effort on the directly rewarded activities and away from the unrewarded activities. The misspecification of an objective performance measurement system thus may result in agents "gaming the system" by optimizing actual instead of intended measures (Baker et al., 1988). As such, contracts based solely on objective performance measures are imperfect and cause incentive distortions. This problem can be mitigated by including (additional) subjective performance assessments (Baker et al., 1994; Höppe & Moers, 2011). Indjejikian & Matejka (2012) involved business unit managers and controllers in a survey study supplemented with in-depth interviews. They find that principals rely more on nonfinancial measures or subjective evaluations in determining local managers' bonuses when local managers have a greater influence on the design of internal accounting systems. This is consistent with principals protecting themselves against agents' asymmetric information or agents' manipulation of objective, accounting measures. Baker et al. (1994) theorize that in some circumstances a combination of objective and subjective measures outperforms an explicit or an implicit contract alone. Moreover, the subjective bonuses provide employees insurance against downside risk in their pay e.g. by filtering out the effect of uncontrollable factors due to interdependencies (Gibbs et al., 2004) or uncertainty (Keeley, 1977; Höppe & Moers, 2011), recalculating incentives when performance targets are too challenging or when the department is facing losses. Subjectivity improves incentive contracting when there is greater trust between the subordinate and the supervisor. This is because the positive effects of subjective bonuses on pay satisfaction and firm performance are larger the longer the supervisor's tenure due to mutual trust (Gibbs et al., 2004). In sum, principals use subjectivity to resolve contracting problems such as incentive distortions (congruity issues) (Höppe & Moers, 2011; Baker et al., 1994), risk concerns (Höppe & Moers, 2011), environmental uncertainty (Keeley, 1977), moral hazard (Cronqvist & Fahlenbrach, 2013; Ke et al., 1999), asymmetric information or agents "gaming" (Indjejikian & Matejka, 2012) or manipulating the system (Gibbs et al., 2004).

However, allowing subjectivity in performance evaluations also has its downsides, as described in the nine articles discussed next. Levin (2003) argues that the use of subjective performance measures necessarily leads to costly disputes and conflicts between the agent and the principal. When agents feel their evaluation is unfair, fairness and conflict concerns will lead to compressed and above average subjective performance evaluations (and thus to higher compensation for the agent) (Golman & Bhatia, 2012; Levin, 2003; MacLeod, 2003). This compressed and above average rating behavior in turn lowers employee performance and firm productivity (Ahn et al., 2010; Golman & Bhatia, 2012). Another concern related to (un)fair rating behavior is favoritism. Subjectivity leads to favoritism where evaluators act on personal preferences toward subordinates to favor some employees over others beyond their true performance. This reduces incentives for the other agents because of increased risk /

uncertainty in evaluations (Prendergast & Topel, 1996). In addition, individuals seem insufficiently aware that a change in the accounting for one subjective measure has spillover effects on the optimal weighting of the other subjective measure in a two-measure incentive system. Consequently, they make performance-measure weighting decisions that are likely to result in misallocations of agent effort (Krishnan et al., 2005). Morse et al. (2011) provide archival evidence that powerful agents are able to shift the weight on performance measures toward the better performing measures. This manipulation practice harms future firm performance. In addition, Prendergast (1993) theorizes that agents have an incentive to conform to what they feel their superior wants to hear. The agent distorts his opinion towards the anticipated opinion of the supervisor. As such, too much weight is put on the opinion of the supervisor, which leads to inefficiencies. Another difficulty in subjective performance measurement is due to principals reneging, i.e. they assess the agent's final performance untruthfully in order to pay less reward to the agent. This is possible because the subjective performance information in the optimal contract is not enforceable (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; MacLeod, 2003; Prendergast & Topel, 1993). This evidently undermines the credibility of subjective performance evaluations in optimal contracts. Therefore, it is important to install mechanisms to enforce the subjective performance measurement. Baker et al. (1994) for example cite that implicit contracts are self-enforcing as principals are concerned with their reputation in the labor market for keeping their promises. Levin (2003) remarks that each party has the option to walk away in a relational contract. To prevent that the principal reneges, the payable reward must not exceed the net present value of the benefits the principal realizes under an ongoing contract. This is the case if the principal's discount rate is small enough. The credibility of optimal contracts with subjectivity added can be considerably improved by restricting subjective incentives to that part of the first-best action that cannot be induced by an explicit contract (Budde, 2007). In addition, "discretionary bonus pools" could prevent the principal from reneging in a situation with multiple agents, because the bonus pool amount is agreed upon ex ante and afterwards the total bonus pool is allocated amongst the agents according to the principal's discretion (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009).

1.3 Discretionary Bonus Pools

In this section, we first discuss a number of theoretical articles on the characteristics and benefits of discretionary bonus pools, after which we discuss the experimental articles challenging the predictions in the theoretical articles. Next, we discuss articles investigating the use of discretionary bonus pools in practice. Table 1.2 follows the structure of this section.

Table 1.2 Discretionary bonus pools

Article	Focus/Results	Research Set-up
Analytical models on discretionary bonus pools		
Baiman & Rajan (1995)	Discretionary bonus pools are an efficient way to incorporate non-contractible information in a two-agent setting.	Analytical model
Baker, Jensen & Murphy (1988)	Free-rider problem associated with ordinary profit-sharing plans: employees receive only approximately 1/n of the increased profits related to their effort (where n is the number of participants in the plan).	Literature review
Rajan & Reichelstein (2006)	When the bonus pool covers many agents and/or the principal's subjective information is precise, discretionary bonus pools are nearly as efficient as explicit contracts.	Analytical model
Rajan & Reichelstein (2009)	In the single-agent case it might be optimal to ignore the subjective signal with discretionary bonus pools.	Analytical model
Experimental papers on discretionary bonus pools		
Bailey, Hecht & Towry (2011)	Managers incorporate noncontractible information to a lesser extent than theoretically expected. Managers who can only allocate a part of the bonus pool incorporate noncontractible information to a greater extent than participants with full discretion.	Experiment (170 business school students)
Fisher, Maines, Peffer & Sprinkle (2005)	Subordinate's performance and compensation is larger when the employer has no discretion over total employee compensation, but discretion over allocation of the compensation pool.	Experiment (237 undergraduate business students)
Maas, van Rinsum & Towry (2012)	Supervisors are more willing to obtain costly performance information on individual agents as it becomes more difficult to distinguish individual contributions to group performance.	Experiment (126 undergraduate business students)
Field studies on discretionary bonus pools		
Gibbs, Merchant, Van der Stede & Vargus (2004)	Discretionary bonuses are used to complement perceived weaknesses in quantitative performance measures and to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998-1999) and 1050 surveys in 326 different dealerships
Ittner, Larcker & Meyer (2003)	Discretion in weighting the measures in a BSC bonus plan led to a focus on quantitative, outcome-oriented financial performance measures that were used in earlier non-discretionary bonus plans.	Field study (a large American retail bank)
Ivancevich (1983)	The more unsatisfactory performing engineers in a team, the more favorable ratings are for satisfactory performing engineers. For scientists no such effect was identified.	Field study with 104 supervisors of 624 engineers and 66 supervisors of 404 scientists working in an American company
Merchant, Chow & Wu (1995)	Incentive plans in Taiwanese and US firms are very similar. In both countries, firms make use of discretionary bonus pools.	Field study (open-ended interviews in 2 US and 2 Taiwanese companies)

Baiman & Rajan (1995) and Rajan & Reichelstein (2006, 2009) theoretically discuss the characteristics and benefits of discretionary bonus pools. As indicated above, subjective performance information is complex and subtle, and therefore difficult to observe and verify by a third party. Since this information is not enforceable (Ahn et al., 2010; Baily et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; MacLeod, 2003;

Prendergast & Topel, 1993), it harms the credibility of subjective performance measurement in optimal contracts. Supervisors can assess a subordinate's final performance untruthfully in order to pay less reward to that subordinate (Bol, 2008). In this context, a discretionary bonus pool is an appealing instrument. The magnitude of the bonus pool is based on an explicit formula agreed-upon ex ante and involving objective performance measures. The entire bonus pool is paid out regardless of the subjective information observed by the supervisor (Baiman & Rajan, 1995). Supervisors cannot change the magnitude of the reward by assessing agent's performance falsely and they have consequently no incentive to do so (Rajan & Reichelstein, 2006). Afterwards, the bonus pool is allocated amongst different subordinates at the supervisor's discretion. Based on subjective information the supervisor can shift a part of the bonus of one subordinate to another, better-performing colleague. Accordingly, a supervisor can use non-contractible information to encourage subordinates (Baiman & Rajan, 1995). The supervisor discretion solves the free-rider problem associated with ordinary profit-sharing plans in large organizations described by Baker et al. (1988). With ordinary profit-sharing plans, employees bear the full cost of exerting effort and yet receive only 1/n of the increased profits (where n is the number of participants in the plan). Discretionary bonus pools take individual effort into account (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009). The use of non-contractible information to motivate subordinates results in a strict Pareto improvement compared to the optimal contract that does not use non-contractible information (Baiman & Rajan, 1995). Discretionary bonus pools are even optimal if a supervisor can only rely on non-verifiable subjective information to create incentives for a group of subordinates. Provided the bonus pool covers a large number of subordinates and/or the supervisors' subjective information is fairly precise, bonus pools based solely on subjective information should be nearly as efficient as explicit contracts based on objective and verifiable information (Rajan & Reichelstein, 2006). When no other subordinate is present, the supervisor incurs an additional cost when the subordinate shirks. In the single-subordinate case it might be optimal to ignore the subjective signal (Rajan & Reichelstein, 2009). The reasoning above explains the popularity of discretionary bonus pools both in practice and in research.

Three experimental articles challenge the predictions made by the theoretical articles discussed above. Fisher et al. (2005) undertook an experiment to examine situations in which the supervisor either has full discretion or no discretion over the *magnitude* of the bonus pool and/or the *allocation* of this bonus pool among subordinates. A compensation scheme in which a supervisor has full discretion to use private information may reduce subordinate opportunism, but allows for supervisor opportunism. The researchers measured the total group output of the subordinates, the bonus allocated to the subordinates and the residual supervisor profit. Both total group output and subordinate compensation appeared to be greater when the supervisor had no discretion over the magnitude of the bonus pool, but discretion over the allocation of the bonus pool. The supervisor's residual profit was higher when he had discretion over the allocation of compensation, while discretion over the magnitude of the bonus pool had no effect on residual profit. So, in general, the discretionary bonus pool outperforms the other experimental situations. This is consistent with Baiman & Rajan (1995). Bailey et al. (2011) experimentally examine situations in which the supervisor has *full or partial* discretion to allocate the bonus pool and/or he is confronted with *positive or negative* noncontractible

information. The findings show that managers incorporate noncontractible information to a lesser extent than theoretically expected by Rajan & Reichelstein (2006) when allocating a bonus pool. When processing performance information, managers in the experiment tended to choose an anchor point and then subsequently adjusted for noncontractible information. This anchoring approach is in contrast to the theoretical approach in Rajan & Reichelstein (2006) in which a manager is supposed to integrate all contractible and noncontractible information into a single, comprehensive performance measure (integrative approach). Managers who use an anchoring approach incorporate noncontractible information into bonus pool allocations to a lesser extent than those who use an integrative approach. In practice, this leads to a reduction in the intended, theoretical benefits of managerial discretion in bonus allocation proclaimed by Rajan & Reichelstein (2006). Participants who can only allocate a part of the bonus pool incorporate noncontractible information to a greater extent than participants with full discretion (Bailey et al., 2011). The third experimental article, by Maas et al., starts from the observation that joint production and unobservability make individual output not readily quantifiable in most jobs (Baker et al., 1988; Levin, 2003; Maas et al., 2012). This impedes the bonus pool allocation discretion of the supervisor in a discretionary bonus pool setting. Maas et al. (2012) investigate the willingness of supervisors to obtain additional, costly information to more accurately assess individual contributions to team output. In their experiment, the aggregate team output is readily available and the individual output can be obtained at an additional cost. The results indicate that supervisors are willing to incur a cost to prevent potential unfairness. Supervisors are more willing to obtain the costly information as it becomes more difficult to distinguish individual contributions to group performance. Additionally, this willingness appeared to be greater for relatively high versus relatively low levels of group performance.

Four articles investigate the functioning of discretionary bonus pools in practice. Ivancevich (1983) provides evidence that a supervisor shifting a part of the bonus from one subordinate to another, better-performing colleague (supervisor allocation discretion) follows a natural reflex. In a field study, Ivancevich (1983) instructed supervisors to evaluate each member of their team individually. Team size ranged from 9 to 44 engineers per team. The supervisors were instructed not to force themselves to come up with distributed evaluations. Despite this instruction, the study shows contrast effects that are very similar to situations where supervisors use allocation discretion in discretionary bonus pools. The more unsatisfactory performing engineers in a team, the more favorable the ratings are for the well performing engineers. Well performing employees are thus the beneficiaries of higher performance ratings and more rewards when unsatisfactory performers are part of the supervisor's team (Ivancevich, 1983). Merchant et al. (1995) investigate the use of discretionary bonus pools in practice. They show that incentive plans in Taiwanese and US firms are very similar in terms of the extent of individual performance-dependent monetary rewards, the extent of group-rewards compared to individual rewards and the amount of subjectivity in evaluations.[8]

Subsequently, Gibbs et al. (2004) find in the car dealership context that discretionary bonuses are used to complement perceived weaknesses in quantitative performance measures

[8] While no cultural differences were found, we should be cautious when interpreting these results as the researchers only investigated 2 US companies and 2 Taiwanese companies (Merchant et al., 1995).

(incompleteness, short-term focus and susceptibility to manipulation) and to provide employees with an insurance against downside risk in their pay (by filtering out uncontrollables due to interdependencies, recalculating incentives when performance targets are too challenging or when department is facing losses). In addition, they find that the use of discretionary bonus pools is positively related to pay satisfaction and firm performance when the manager has long tenure. Finally, Ittner et al. (2003) undertake a field study on the introduction of a subjective BSC-based bonus plan containing six categories of financial and nonfinancial performance measures in a large American retail bank. The supervisor could subjectively decide on the weighting of the different performance measures. Ittner et al. (2003) were confronted with a number of downsides of this practice. The discretion in weighting the measures in the bonus plan led the supervisors to ignore many performance measures, to change weightings from period to period and to include factors that were not even performance measures, although this was not allowed. In other words, quantitative, outcome-oriented financial performance measures as used in earlier non-discretionary bonus plans remain dominant. The high level of discretion and the related uncertainty in the criteria used for bonus determination made many subordinates complain about favoritism. Afterwards, the firm chose for a non-discretionary, formulaic bonus plan based solely on revenues. This field study points out that psychology may be more important in explaining firm's measurement practices than optimal contracting.

1.4 Judgment Biases & Debiasing

Rating inaccuracy caused by performance evaluation biases is perceived as one of the main problems of introducing subjectivity into compensation contracts (Bol, 2011). Supervisors need to invest time and effort in gathering accurate information on employee performance (Bol, 2011) and are not the residual claimants of subordinates' output, which leaves room for supervisors' preferences (Prendergast & Topel, 1993). This section provides an overview of different judgment biases by supervisors. These judgment biases may impede or reinforce the proclaimed benefits of subjectivity in performance measurement discussed in the previous section on optimal contracting. Table 1.3 provides an overview of the articles on judgment biases and debiasing. Debiasing concerns practices to resolve judgment biases. It lists the articles discussing compression bias or centrality bias, the articles focusing on biases related to the BSC, the articles concerning biases related to personal characteristics, and a number of articles about biases related to accompanying or competitive information.

Table 1.3 Judgment biases and debiasing in subjective performance evaluations

Article	Focus/Results	Research Set-up
Leniency bias and compression bias		
Ahn, Hwang & Kim (2010)	Subjective measures provide less incentive than objective measures because of the lack of variation in scores (compression bias).	Archival (13 government-invested companies, Republic of Korea, 1990-2006)
Baker, Jensen & Murphy (1988)	Biased and inaccurate performance evaluations reduce effectiveness of incentives and productivity.	Literature review
Bol (2011)	Information-gathering costs and strong subordinate-supervisor relationships increase centrality bias and leniency bias. Centrality bias decreases performance improvement and leniency bias increases future performance.	Archival study (5 branch offices of a Dutch financial service provider, 2003-2004, 198 employees
Duarte, Goodson & Klich (1994)	Subjective performance ratings in high-quality supervisor-subordinate relationships are always high. Ratings in low-quality relationships are consistent with objective performance measures in the short run, but high in the long run.	Questionnaire (261 supervisor-subordinate pairs in an American telephone company)
Golman & Bhatia (2012)	Noise in the performance signal and a stronger aversion to unfairly low ratings than to overly high ones result in leniency bias. Noise in the performance signal results in compression bias. Both biases hurt agent's performance.	Analytical model
Judge & Ferris (1993)	Greater opportunity to observe a subordinate's job performance resulted in higher performance ratings.	Questionnaire (81 nurses and their supervisors in an American hospital)
Kane, Bernardin, Villanova & Peyrefitte (1995)	Leniency is a relatively stable response tendency by individual raters.	3 field studies: one: 328 patrol officers, 38 sergeants and 14 lieutenants in a police department, two: 243 nurses and 31 head nurses, three: 44 supervisors of 376 social workers
Levin (2003)	Self-enforced relational contracts with moral hazard result in compression bias.	Analytical model
MacLeod (2003)	Optimal contracts with subjective evaluations can result in compression bias and leniency bias.	Analytical model
Moers (2005)	The use of multiple objective performance measures and the use of subjective performance measures are related to compression and leniency bias.	Archival study (124 subordinates in a Dutch, maritime industrial firm, 1998)
Prendergast & Topel (1993)	Supervisors are not the residual claimants of subordinates' output, which makes supervisors' preferences and biases such as leniency bias, favoritism and compression bias possible.	Literature review
Biases observed in a BSC context		
Banker, Chang & Pizzini (2004)	Evaluators focus more on common measures than on unique measures. Evaluators focus more on strategically linked measures than non-linked measures only when evaluators are provided with detailed information about BU-strategy.	Experiment (480 MBA students)
Cardinaels & van Veen-Dirks (2010)	When there are performance differences in the financial performance measures, evaluators that use a BSC-format place more weight on the financial performance measures than evaluators using an unformatted scorecard. When there are performance differences in the non-financial performance measures, evaluators evaluate similarly in both formats.	2 experiments (144 business program students)

Article	Focus/Results	Research Set-up
Choi, Hecht & Tayler (2012)	Surrogation: managers forget that performance measures are imperfect representations of the underlying strategic construct.	Experiment (79 graduate business students)
Ding & Beaulieu (2011)	Participants who were induced to feel good (bad) gave higher (lower) evaluation scores to divisional managers.	Experiment 1 (104 MBA students) and Experiment 2 (32 MBA students)
Humphreys & Trotman (2011)	When *all* the performance measures are strategically linked, but no strategy information is provided, common measure bias exists. When strategy information is present and *all* performance measures are strategically linked, then common measure bias disappears.	Experiment (92 executive MBA students)
Libby, Salterio & Webb (2004)	Either the requirement to justify a performance evaluation to a superior or improving perceived quality of the BSC measures via an independent third-party report decreases common measure bias.	Experiment (227 MBA students)
Lipe & Salterio (2000)	Superiors use only the common performance measures to evaluate performance of the business unit in a BSC-context.	Experiment (58 MBA students)
Lipe & Salterio (2002)	Performance evaluations are affected by organizing the measures into the BSC categories when multiple below-target (or above-target) measures are contained within a category but those evaluations are not affected when the above/below-target measures are distributed across the scorecard's four categories or when the same measures are presented without the organizing BSC categories.	Experiment (78 MBA students)
Tayler (2010)	Managers who are involved in selecting strategic initiatives perceive those initiatives afterwards as more successful than managers who are not involved in the initiative-selection process. Simply framing the scorecard as a causal chain is not sufficient to mitigate these effects, but framing the scorecard as a causal chain *and* involving managers in the selection of scorecard measures, mitigates the effects.	Experiment (132 MBA students)
Biases related to personal characteristics		
Biernat & Sesko (2013)	Evaluations of mixed-sex work teams' performance: women were solely judged lower in a white pair work team. Black women were not affected by gender bias.	2 experiments (142 and 283 undergraduate students respectively)
Elvira & Town (2001)	Racial differences between subordinate and supervisor lead to lower ratings for both black and white subordinates.	Field study (316 salespersons in a large, American company)
Judge & Ferris (1993)	Demographic similarity and the supervisor-subordinate relationship significantly influenced performance rating.	Questionnaire (81 nurses and their supervisors in an American hospital)
Pulakos & Wexley (1983)	Supervisors appraise dissimilar subordinates significantly lower.	Questionnaire (171 supervisor-subordinate relationships in manufacturing, retailing, government and service organizations)
Varma & Stroh (2001)	Both male and female supervisors rate subordinates of the same sex higher.	220 surveys of supervisors in the communications industry
Wayne & Liden (1995)	Demographic similarity and subordinates' impression management influence performance ratings.	Survey (111 supervisor-subordinate pairs in nonacademic jobs at 2 American universities)

Article	Focus/Results	Research Set-up
Biases related to accompanying or competitive information		
Bol & Smith (2011)	Supervisors bias their subjective evaluations of performance to be consistent with an accompanying objective performance measure.	Experiment (216 supervising employees at an university)
Dossett & Greenberg (1981)	A worker who initially suggested a high goal received a significantly higher performance score than a worker who suggested a low goal.	Experiment with 80 undergraduate students
Ghosh & Lusch (2000)	Unfavorable outcomes negatively influence subjective performance evaluations.	Archival study in 204 stores of an American retailer
Heneman & Wexley (1983)	Performance ratings are less accurate when rating is delayed and when only a small amount of information is observed.	Experiment (180 undergraduate business students)
Hogan (1987)	Ratings will be lower when a subordinate's actual performance disappoints a supervisor's expectations about that performance.	Questionnaire (49 subordinate-supervisor pairs in an American bank)
Ittner, Larcker & Meyer (2003)	Supervisors focus on quantitative, outcome-oriented financial performance measures.	Field study (a large American retail bank)
Jacobs & Kozlowski (1985)	As raters have more opportunity to observe ratee behavior, the magnitude of halo error increases.	3 consecutive ratings (1031, 976 and 876 students respectively).
Tan & Jamal (2001)	Average superiors evaluate work done by outstanding subordinates more favorably than work done by average subordinates when they know the identity of the work preparer. Outstanding superiors are not affected by the halo effect.	Experiment (40 audit seniors and 20 audit managers)

Compression bias or centrality bias refers to the tendency to compress performance ratings, which results in less variance in ratings than justified by the variance in actual performance. Leniency bias is the tendency to inflate subordinate's performance rating such that subordinate's performance is assessed to be above average (Baker et al., 1988; Bol, 2011). These important forms of performance evaluation bias have received quite some research attention. According to the theoretical articles on optimal contracting, supervisor's subjective performance evaluations will be compressed and rated above average if the supervisor and subordinate disagree upon the subjective performance evaluation or when moral hazard is present (Levin, 2003; MacLeod, 2003). This is due to the tradeoff between reducing the cost of conflict between subordinate and supervisor ex post (Bol, 2011; MacLeod, 2003) and providing incentives to the subordinate ex ante (MacLeod, 2003). Supervisors need to invest time and effort in gathering accurate information on employee performance (Bol, 2011). They bear all of the monitoring costs but receive little of the benefit from conducting more accurate evaluations (Baker et al., 1988). Whenever information-gathering costs increase, they invest less time and effort in gathering accurate information on employee performance. Therefore, it is not surprising that empirical evidence indicates that centrality bias and leniency bias are positively related to information-gathering costs (Bol, 2011), to the use of multiple objective performance measures and to the use of subjective performance measures (Moers, 2005). Additionally, uncertainty about subordinate performance leads to compressed ratings. The analytical model by Golman & Bhatia (2012) indicates that when a supervisor is uncertain about subordinate performance and he has a stronger aversion to unfairly low ratings than to

overly high ratings (due to fairness or conflict concerns), he will inflate performance ratings according to his preferences. Kane et al. (1995) provide empirical evidence that inflating performance ratings is a relatively stable rater tendency.

In contrast to the reasoning above, Duarte et al. (1994) and Judge & Ferris (1993) provide empirical evidence about the presence of leniency in a different way: they found that a greater opportunity to observe a subordinate's job performance actually resulted in higher performance ratings. Along this reasoning, the supervisor-subordinate relationship significantly influences supervisors' affection toward subordinates and consequently influences performance ratings indirectly through supervisors' affection (Judge & Ferris, 1993). Strong subordinate-supervisor relationships increase centrality bias and leniency bias (Bol, 2011). Both in the short run and the long run, subjective performance ratings in high-quality supervisor-subordinate relationships are high, regardless of objective performance measures. Ratings in low-quality relationships are consistent with objective performance measures in the short run, but high in the long run, regardless of objective measures (Duarte et al., 1994).

The literature shows mixed results on the effect of centrality bias and leniency bias on subordinate performance. Ahn et al. (2010) examine the effect of discriminability (variation in performance scores) on subordinate performance empirically. Their findings show that subordinate performance improvement increases with the degree of discriminability. Subjective measures provide less incentive than objective measures because of the lack of discriminability (compression bias). This results in a decrease in performance improvement (Ahn et al., 2010; Bol, 2011). Biased and inaccurate performance evaluations reduce productivity by reducing effectiveness of incentives (Baker et al., 1988). In other words, leniency bias and centrality bias hurt the agent's performance (Golman & Bhatia, 2012). In contrast, Bol (2011) reveals that leniency bias increases future performance due to increased perceived fairness of the incentive system.

In the following paragraphs, we discuss a number of experiments investigating judgment biases in the context of the BSC or another strategic performance evaluation framework. The BSC is a framework devised by Kaplan & Norton (1992) containing a large set of performance measures that capture the drivers of the firm's desired business strategy along four categories: financial performance, customer relations, internal business processes and the organization's learning and growth activities. The distinctive feature of the BSC is that the performance measures are linked with each other, and that they express cause-and-effect relationships that lead to the implementation of the intended strategy. The BSC can be used to evaluate the performance of a business unit or a business unit manager (Lipe & Salterio, 2000, 2002). Lipe & Salterio (2000) point out that BSCs include some performance measures common to multiple business units and other performance measures that are unique to a particular business unit. Based on an experiment, they discover that supervisors make only use of the common measures when evaluating the performance of different business units i.e. common measure bias is present (Lipe & Salterio, 2000). Banker et al. (2004) confirm this result in an experiment in which some measures are strategically linked and others are not: evaluators focus more on common performance measures than on unique measures. However, this result does not hold when a number of the BSC performance measures are strategically linked *and* detailed information on the strategic linkages is provided. In that case, evaluators focus more on

strategically linked unique measures than on non-linked measures that are common (Banker et al., 2004). In contrast, when strategy information is provided to managers *and* only *some* measures are strategically linked, common linked measures get more attention than unique linked measures (Banker et al., 2004; Humphreys & Trotman, 2011). Humphreys & Trotman (2011) further experimentally demonstrate that common measure bias exists when only some or *all* the performance measures are strategically linked, but no strategy information is provided. However, when strategy information is present and *all* performance measures are strategically linked, common measure bias disappears, but Libby et al. (2004) propose two methods to overcome common measure bias. The first one is to introduce a requirement to justify the performance evaluation to a superior. An alternative is to improve the perceived quality of the BSC performance measures via the provision of an independent third-party report (Libby et al., 2004).

Other experimental studies focus on how BSC framing affects performance evaluations. Lipe & Salterio (2002) find that when multiple below-target (or above-target) measures are contained within a single BSC category, performance evaluations are different from performance evaluations using the same measures, but without the BSC categories framework. However, when the above/below-target performance measures are distributed across the four categories of the BSC, evaluations are not different from evaluations using these same measures, but without the BSC categories framework. The reason is that when performance on measures within a group is consistent (e.g. consistently above-target), the decision maker may perceive that the measures are related. Consequently, he reduces the impact of the individual performance measures on his or her judgment. In contrast, when the same measures are presented without the organizing BSC categories (or are scattered across BSC categories), the perception of relations among these measures and the resulting reduction in decision weights are less likely. Also Cardinaels & van Veen-Dirks (2010) investigate the effect of the presentation of performance measures on performance evaluations of two business unit managers, especially how evaluators weight financial and non-financial measures. When the performance difference between the two managers is located in performance measures in the financial category, evaluators that use a BSC-format place more weight on the financial category performance measures than evaluators using an unformatted scorecard. In contrast, when the performance difference is located in performance measures of one of the three non-financial categories, the weight placed upon these non-financial category measures is similar for the BSC-format and the unformatted scorecard. In a subsequent experiment, Cardinaels & van Veen-Dirks (2010) use performance markers to indicate above- target, on-target or below-target performance. In this setting, evaluators that use a BSC-format weight financial and non-financial performance differences more heavily than evaluators using an unformatted scorecard. Managers who are involved in selecting strategic initiatives perceive those initiatives afterwards as having been more successful than managers who are not involved in the initiative-selection process. Simply framing the scorecard as a causal chain is not sufficient to mitigate these effects, but framing the scorecard as a causal chain *and* involving managers in the selection of scorecard measures, mitigates the effects of manager involvement in initiative selection on initiative performance evaluation (Tayler, 2010).

Two more studies undertaken in a BSC context investigate how incentive payment affects the evaluation of performance by managers. Choi et al. (2012) remark that firms develop strategic performance measurement systems (SPMS) that translate strategy into imperfect performance measures of the true strategic construct. Unfortunately, managers fail to acknowledge that performance measures are imperfect representations of the strategic construct and act as if the performance measures are the constructs of interest (surrogation). Surrogation is increased by incentive compensation. This effect is larger when compensation is based on a single measure of the strategic construct compared to when it is based on multiple measures of a strategic construct (Choi et al., 2012). Ding & Beaulieu (2011) show that participants who were induced to feel good (bad) gave higher (lower) evaluation scores to divisional managers both in a setting with only two performance measures and in a setting with a 16-measure BSC. Financial incentives eliminated the mood congruency bias in the two-performance-measure- condition and in the condition with a simplified BSC with only eight measures, but not in the 16-measure-BSC-condition. Financial incentives thus can reduce the bias if the BSC is not too extensive (Ding & Beaulieu, 2011).

The literature links judgment biases also to personal characteristics of the subordinate and/or the supervisor. Social and situational influences are important in the performance-rating process. Demographic similarity and the supervisor-subordinate relationship significantly influence supervisors' affection for subordinates and influence performance rating indirectly through supervisors' affection (Judge & Ferris, 1993; Wayne & Liden, 1995). Supervisors appraise the performance of subordinates whom they perceive as being dissimilar to themselves significantly lower (Pulakos & Wexley, 1983) e.g. after controlling for performance, racial differences between subordinate and supervisor lead to lower ratings for both black and white subordinates (Elvira & Town, 2001) and both male and female supervisors rate subordinates of the same sex higher (Varma & Stroh, 2001). Black subordinates get lower subjective performance ratings than whites (Elvira & Town, 2001). Biernat & Sesko (2013) investigates the evaluations of mixed-sex work teams' performance after having performed a masculine task. A mixed-sex team consists of a white pair, a black pair or a mixed-race pair. Women's competence was solely judged lower in a white pair work team. Black women were not affected by gender bias (Biernat & Sesko, 2013).

Finally, this section discusses the literature on biases related to accompanying or competitive information about the subordinate's performance, i.e. the outcome effect, the assimilation/spillover effect and the halo effect. The outcome effect captures that when an objective outcome measure is positive (negative), supervisors tend to evaluate the subordinate positively (negatively), regardless of the actual appropriateness of the decision resulting in that outcome. For instance, Ghosh & Lusch (2000) document how subjective performance evaluations of store managers were negatively influenced by unfavorable objective outcome knowledge. Similarly, Ittner et al. (2003) document how supervisors tend to focus on outcome-oriented financial performance measures when evaluating subordinates. Taking outcomes into account that do not reflect subordinates' performance will affect not only the quality of the subjective performance evaluation but will also incorrectly reward/penalize subordinates (Ghosh & Lusch, 2000). The assimilation or spillover effect is a bias very similar to the outcome effect. In this case, supervisors bias their subjective evaluations of performance on

one dimension to be consistent with an objective measure of performance on a separate and unrelated dimension (Bol & Smith, 2011). Likewise, Duarte et al. (1994) document that ratings are consistent with objective performance measures, however only with low-quality relationships in the short run. The halo effect relates to the observation that the supervisor's prior expectations about a subordinate's performance have an effect on later ratings of that performance. When a subordinate's actual performance disappoints a supervisor's expectations about that performance, subsequent ratings will be lower, regardless whether the actual performance is better or worse than expected. The more often supervisors must use objective measures like rating formats or strict procedures, the more likely ratings are to be accurate (Hogan, 1987). A supervisor's general impression formed from prior interactions may impair the supervisor's ability to objectively assess the subordinate's current work (Tan & Jamal, 2001). More specifically, Tan & Jamal (2001) show in an audit-context that average superiors evaluate work done by outstanding subordinates more favorably than work done by average subordinates when they know the identity of the work preparer, but not when the identity of the work preparer is unknown. Outstanding superiors are not affected by the perceived competence of the subordinate preparing the work. Dossett & Greenberg (1981) investigate how employees can steer this halo effect. They examine the effect of who sets the performance goal (self-set, participative or assigned) and performance outcome on employee's performance evaluation. Their findings indicate that a worker who initially suggested a high goal received a significantly higher performance score than a worker who suggested a low goal, consistent with the halo effect. Besides, as raters have more opportunity to observe ratee behavior (i.e. higher familiarity between supervisor and subordinate), the magnitude of halo error increases (Jacobs & Kozlowski, 1985). Performance ratings are less accurate when rating is delayed (instead of immediate rating) and when only a small amount of information is observed (Heneman & Wexley, 1983).

1.5 Perceived Fairness

Whether a performance evaluation is perceived as 'fair' depends amongst others on influence activities (i.e. attempts of subordinates to influence the evaluation of the supervisor) (Du et al., 2012), favoritism (i.e. supervisors acting on personal preferences toward subordinates to favor some subordinates over others beyond their true performance) (Du et al., 2012; Prendergast & Topel, 1996), procedural justice (i.e. the perceived fairness of the means and procedures used to determine the subjective performance evaluations) (Dulebohn & Ferris, 1999; McFarlin & Sweeney, 1992) and whether supervisors adjust their subjective performance evaluations when uncontrollable factors affect subordinates results (Bol & Smith, 2011; Ghosh & Lusch, 2000). Perceived fairness is of utmost importance in subjective performance evaluations. Table 1.4 summarizes the articles dealing with perceived fairness in subjective performance evaluations. The majority of these articles deal with influence activities, favoritism and procedural justice, while some also deal with controllability.

Table 1.4 Perceived fairness in subjective performance evaluations

Article	Focus/Results	Research Set-up
Influence activities, favoritism and procedural justice		
Burney, Henle & Widener (2009)	The higher organizational justice, the higher employee performance.	Survey (242 persons, in 47 branches of a large American financial services organization
Du, Tang & Young (2012)	Both influence activities and government favoritism affect the evaluation positively.	Archival study (63 state-owned enterprises (SOEs), 2005-2007) and interviews (6 CFOs of SOEs)
Dulebohn & Ferris (1999)	Supervisor-focused influence tactics are associated with positive procedural justice evaluations, but job-focused influence tactics were associated with negative evaluations.	Field study (128 subordinates and 23 supervisors in a food service department)
Hartmann & Slapnicar (2009)	Subordinate's trust in the superior depends on the formality of the performance evaluation procedure.	Survey (160 departmental managers in 11 Slovenian commercial banks)
Hartmann & Slapnicar (2012)	The amount of subordinate's voice in the performance evaluation process has a positive effect on justice perceptions.	Survey (178 upper-middle managers in 12 Slovenian commercial banks)
Hartmann, Naranjo-gil & Perego (2010)	Initiating structure-leaders and consideration-leaders both enhance evaluation fairness in their own way.	Survey (196 middle-level managers in 11 Dutch organizations)
McFarlin & Sweeney (1992)	Procedural justice is an important predictor of organizational commitment.	Survey (675 employees of an American bank)
Prendergast & Topel (1996)	Favoritism causes firms to use bureaucratic rules in pay decisions and firms place too little weight on supervisor appraisals. Favoritism reduces incentives because of increased risk in evaluations.	Analytical model
Prendergast (1993)	Subordinates have an incentive to conform to what they feel their superiors want to hear.	Analytical model
Wayne & Liden (1995)	Demographic similarity and subordinates' impression management influence performance ratings.	Survey (111 supervisor-subordinate pairs in nonacademic jobs at 2 American universities)
Woods (2012)	Supervisors use downward adjustments to performance evaluations to encourage the departure of certain subordinates.	Field study (272 observations and 66 surveys in an internal audit organization in 2006)
Controllability		
Bol & Smith (2011)	Supervisors adjust their evaluations when an uncontrollable factor decreases the subordinate's objective measure, but they do not adjust the evaluations when the uncontrollable factor increases subordinate's objective measure.	Experiment (216 non-academic supervising employees at a university)
Ghosh & Lusch (2000)	(Un)controllable outcomes (do not) influence performance evaluations, but central management determinants of outcome, which are uncontrollable , influence evaluations.	Archival study in 204 stores of an American retailer
Gibbs, Merchant, Van der Stede & Vargus (2004)	Subjective bonuses are used to provide employees insurance against downside risk in their pay.	Archival study (526 department managers in 250 American car dealerships in 1998-1999) and 1050 surveys in 326 different dealerships
Giraud, Langevin & Mendoza (2008)	For uncontrollable factors external to the company, managers do not prefer the controllability principle, but for internal, uncontrollable factors, managers prefer the principle.	Survey (265 French managers)

Article	Focus/Results	Research Set-up
Govindarajan (1984)	Superiors of BUs with higher environmental uncertainty will use a more subjective performance appraisal and superiors of BUs with lower environmental uncertainty will use a more formula-based performance evaluation.	Interviews (managers of business units within 8 Fortune 500 firms and 58 surveys)
Höppe & Moers (2011)	Discretionary bonuses are used for risk-reduction purposes.	Archival study (1,753 firm-year-observations for 424 American, publicly listed firms, 1998-2002)
Wong-On-Wing, Guo, Li & Yang (2007)	Top managers do not automatically take into account the quality of strategy (uncontrollable factor for divisional managers) in performance evaluation of divisional units using BSC. Divisional managers automatically consider the quality of strategy without being asked to do so.	Experiment (68 MBA students)

Subjectivity leads to favoritism when evaluators act on personal preferences toward subordinates to favor some employees over others beyond their true performance (Prendergast & Topel, 1996). When other employees discover the favoritism, it leads to a decrease in procedural justice, and associated with that, it results in a decrease in employees' motivation and organizational commitment (McFarlin & Sweeney, 1992; Prendergast & Topel, 1996).

Woods (2012) even provides evidence of how supervisors appear to use downward performance adjustments in order to encourage the departure of certain subordinates. As such, subordinates have an incentive to conform to what they feel their superiors want to hear (Prendergast, 1993). In order to constrain favoritism and to induce subordinates to report information honestly, analytic results show that firms will deemphasize incentive pay for subordinates, increase the use of bureaucratic rules in pay decision and place too little weight on supervisor appraisals, giving too much weight to noncorruptible, objective measures such as seniority (Prendergast, 1993; Prendergast & Topel, 1996). Employees' perceptions of justice in an organizational context increase with the degree to which employees think that the strategic performance measurement system (SPMS) reflects a strategic causal model, the degree to which the SPMS is technically valid[9] (Burney et al., 2009) and the amount of subordinate's voice in the performance evaluation process (Hartmann & Slapnicar, 2012). Subordinate's trust in the superior depends on the formality of the performance evaluation procedure (i.e. procedural justice, represented by explicit targets, clear metrics and clear bonus allocation rules) as well. This is because formality increases the perceived quality of feedback and perceptions of procedural justice. Formality matters more for trust formation to those managers that are in functions with less contractible outputs (Hartmann & Slapnicar, 2009). Supervisors that score high on consideration leadership style, i.e. supervisors concerned with the promotion of subordinates' well-being through supportive and pleasant relationships, significantly affect procedural fairness directly. Supervisors high on initiating structure leadership style, i.e. supervisors clearly indicating the roles of their subordinates toward the attainment of organizational goals, by in detail deciding what will be done and how it should

[9] A technically valid SPMS provides employees with performance measures information that is accurate, accessible, understandable, reliable and timely. Employees have access to the performance measures information, understand what it means and how to use it in carrying out their job. If this information is used to define employee's incentive compensation, employees will likely think that their evaluations accurately reflect their effort and their expectations (Burney et al., 2009).

be done, are effective in enhancing fairness towards their subordinates by clearly communicating expectations and setting objective standards in performance evaluation procedures (Hartman et al., 2010). Burney et al. (2009) show that firms do not necessarily need to introduce subjectivity into the incentive contracts to improve performance. If firms clearly communicate the characteristics of the SPMS incentive plan such that employees perceive this plan has a high degree of technical validity and it highly reflects the organization's strategic causal model, this will enhance employees' perceptions of justice in an organizational context, which will in turn affect employees' performance positively as well (Burney et al., 2009). Evidence in the context of a Chinese government agency evaluating the performance of different state-owned enterprises reveals that both influence activities and favoritism affect performance evaluation positively (Du et al., 2012). Wayne & Liden (1995) develop and test a theoretical model to understand the effect of subordinate impression management[10] (i.e. influence activities) on supervisor performance ratings. They find that demographic similarity and subordinates' impression management influence performance ratings through supervisors' liking of and perceived similarity to subordinates (Wayne & Liden, 1995). Dulebohn & Ferris (1999) investigate the impact of employees' use of influence tactics on their evaluations of the fairness of the performance evaluation process. They distinguish between two categories of influence tactics: supervisor-focused tactics, tactics used by employees to be better liked by their supervisors such as flattery and doing favors, and job-focused tactics, tactics used to self-promote and appear competent. Supervisor-focused influence tactics are associated with positive employee evaluations of procedural justice, but job-focused influence tactics were associated with negative employee evaluations of procedural justice.

The literature on optimal contracting (see section 2) revealed that fairness considerations of supervisors lead to optimal contracts whereby supervisor's subjective performance evaluations are compressed and above average (compression bias and leniency bias) (MacLeod, 2003; Golman & Bhatia, 2012). Bol (2011) empirically investigates the effect of leniency bias on future performance in a financial service provider. She shows that leniency bias increases future performance. This is explained by the fact that subordinates overestimate their abilities relative to their supervisors. As such, leniency bias results in an increase in congruence between the rating the employee thinks to deserve and the rating the subordinate actually receives. Lenient ratings are more in line with the expectations of self-over-estimating employees and consequently improve perceived fairness of the incentive system and, in turn, employee motivation (Bol, 2011). Moreover, empirical research in a retailer (Ghosh & Lusch, 2000), publicly-listed companies (Höppe & Moers, 2011) and car dealerships (Gibbs et al., 2004) indicates that supervisors take factors uncontrollable to the subordinate but affecting subordinate's performance into account when evaluating subordinate's performance (Ghosh & Lusch, 2000; Höppe & Moers, 2011; Gibbs et al., 2004). Research in eight Fortune 500 firms reveals that superiors of business units which face higher environmental uncertainty will use a more subjective performance appraisal approach and superiors of business units which face

[10] Impression management are those behaviors individuals employ to protect their self-images, influence the way they are perceived by significant others, or both. The individuals try to look more appealing/favorable to their superior or to peers. This can be accomplished with smiling, eye contact, touching, verbally agreeing, flattery, favor-doing, opinion conformity with the superior or the peer etc. (Wayne & Liden, 1995).

lower environmental uncertainty will use a more formula-based performance evaluation approach (Govindarajan, 1984). In the experimental setting of Bol & Smith (2011), supervisors adjust their subjective performance evaluations when an uncontrollable factor decreases the subordinate's unrelated, objective measure (i.e. they compensate for bad luck), but they do not adjust the evaluations when the uncontrollable factor increases subordinate's objective measure (i.e. they do not punish for good luck). This is consistent with fairness considerations (Bol & Smith, 2011). Supervisors provide employees insurance against downside risk in their pay (filtering out uncontrollable factors due to interdependencies, recalculating incentives when performance targets are too challenging or when department is facing losses) (Gibbs et al., 2004). In their field study, Ghosh & Lusch (2000) document how outcome determinants over which subordinates have control influence their subjective performance evaluations and environmental determinants of outcome over which they have no control do not influence their evaluations. However, inconsistent with the reasoning above, Ghosh & Lusch (2000) find that determinants of outcome decided by central management also influence subordinate's performance evaluations, although they have no influence over those determinants. Wong-On-Wing et al. (2007) find a similar result: they find that supervisors do not take into account the quality of strategy, an uncontrollable factor for their subordinates, when evaluating the subordinates unless they are explicitly required to do so. In contrast, subordinates automatically consider the effect of the quality of strategy on their performance without being prompted to do so. When the uncontrollable factors are internal (interdependencies due to decisions made by colleagues or superiors), managers prefer that their supervisors take these uncontrollable factors into account. Yet, at the same time, when the uncontrollable factors are external to the company, managers do not want their supervisors to adjust for these uncontrollable factors. They fear that the procedure to neutralize the effect of the uncontrollable factors may result in subjectivity in an unfair way (Giraud et al., 2008).

1.6 Avenues for Future Research

Based on our literature review, we can identify several avenues for future research. The structure of this section follows the structure of this paper. We start with possible extensions related to optimal contracting. Afterwards we discuss avenues for future research dealing with discretionary bonus pools and we end with sections discussing research opportunities related to 'biases and debiasing' and 'perceived fairness' respectively.

The current state of the literature indicates a number of research opportunities in optimal contracting with subjective performance evaluations. First, older studies such as Gibbs et al. (2004) or Govindarajan (1984) discuss subjectivity in a very general way. These studies just state 'the use of subjectivity' in a broad, general sense without distinguishing between the different forms of subjectivity such as subjective weightings of objective performance measures, the use of subjective performance measures, a subjective performance evaluation by the supervisor or the possibility to take into account non-prespecified factors in the performance evaluation ex post. This research area would benefit from acknowledging that many different types of subjectivity exist. Researchers can extend the classification of different types of subjectivity and investigate the use of more specific types of subjectivity to broaden our understanding of the use, costs, benefits and consequences of different types of subjectivity

(Bol, 2008; Du et al., 2012; Höppe & Moers, 2011; Ittner et al., 2003; Ke et al., 1999; Woods, 2012). Second, subjectivity does not occur in isolation; firms make use of both objective and subjective performance measures. Nonetheless, many studies do not take the total incentive contract into account when examining subjectivity. The relationships among the various compensation package elements and the specific situations, in which the reward packages are used, could be investigated (Gibbs et al., 2004). For example, future research could examine whether subjective and objective performance measurement act as complements or substitutes (Bol, 2008). How does rating behavior or the optimal contract of the supervisor differ when both objective and subjective performance measures are used? Research, thus far, has not been able to answer these questions. Third, all agency theoretical models described earlier in this paper are quite simplistic. They can be adapted such that they describe a more realistic setting. E.g., MacLeod (2003) deals with a risk-neutral principal and a risk-averse agent and Budde (2007) only investigates the case with a risk neutral principal and agent. The models of MacLeod (2003) and Budde (2007) would be improved if the principal could be risk-averse as well or when principals and agents could transfer risk from one to the other by making use of their private information (MacLeod, 2003). Also, these models could be extended to the case with multiple subjective evaluators (Baker et al., 1994). Finally, the agency literature has focused extensively on the determinants of optimal weights from the supervisor perspective, but has paid little attention to the implications of those optimal weights on subordinate motivation (Ahn et al., 2010).

Future research can also build on the literature on discretionary bonus pools. According to agency theory, a principal/supervisor designs a contract that maximizes firm value. However, most firms are multi-layered and the principal/supervisor designing the contract is not the residual claimant. As such, he has little incentive to aim for the optimal, value-maximizing contract. Instead, to some extent personal preferences will introduce subjectivity in the contract design phase (Baker et al., 1988; Rajan & Reichelstein, 2009). In addition, the discretionary bonus pool optimal contracting literature assumes that the principal agrees ex-ante to optimally allocating the bonus pool and ex post he has no incentive to do otherwise. However, in real-life the principal might not make the optimal bonus pool allocation due to favoritism, influence activities, collusion among agents or sabotage of the performance of one agent by another agent, which distorts the performance information of that other agent (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006, 2009). Future research could explore the conditions under which these implementation problems arise. For example, future research could investigate what effect the supervisor, being the residual claimant or not, has on the structure of the optimal contract.

In general, researchers could further examine how incentive structures of both principal and agent affect research outcomes (Maas et al., 2012). Research focusing on the circumstances under which we would expect to find positive and negative consequences of performance measurement systems on managerial performance would be beneficial (Franco-Santos et al., 2012). Researchers could also investigate employees' responses (i.e., effort and／or performance) to managers' use of allocation discretion. Employees may be proactive in providing favorable noncontractible information to their superiors, and such behavior may vary across different levels of discretion and／or given their perceptions of managers' allocation

processes (Bailey et al., 2011). Additionally, present research focuses on the performance of discretionary bonus pools in a single period. These studies can be extended to a setting with repeated interaction over multiple periods (Rajan & Reichelstein, 2006). In this context supervisors' reputation of trustworthiness might become relevant. Next, we discuss future research opportunities related to the 'biasing & debiasing' stream. In general, there is a need for further research into the role of social and contextual factors in the performance evaluation process (Duarte et al., 1994; Dulebohn & Ferris, 1999; Franco-Santos et al., 2012). The question whether superiors' evaluative behavior could be explained in terms of the context remains unanswered (Hartman et al., 2010). Judge & Ferris (1993) consider the effects of several key aspects of social contexts, such as supervisor-subordinate demographic similarity, supervisor-subordinate work relationship, supervisor's span of control, supervisor's experience, supervisor's affection for the subordinate and supervisor's opportunity to observe subordinate's job performance, on the performance-ratings process. Future research could expand the variables studied and provide a deeper assessment of the causal relationships among those variables, and thus a more informed understanding of the performance-rating process (Judge & Ferris, 1993). Especially research integrating a full range of social context variables with the cognitive processes of supervisors in observing, storing, and recalling data about subordinates is lacking. How do situational variables such as organizational level, work group size, technology and task interdependence influence the processing of information (Wayne & Liden, 1995)? Future research could also take a look at how different levels of the time delay of the rating and/or of the amount of information provided, affect performance rating accuracy (Heneman & Wexley, 1983). Research could investigate whether a particular bias is influenced by person-specific characteristics (Kane et al., 1995; Moers, 2005; Maas et al., 2012; Pulakos & Wexley, 1983) or social context factors (Duarte et al., 1994; Wayne & Liden, 1995), or both. Both internal factors like cognitive consistency, ego enhancement, commitment, and external ones like organization systems, rewards, and social pressures may explain judgment biases (Hogan, 1987). Given equal circumstances, differences in the tendency to rate could reflect personality or information-processing differences among supervisors. Additional research may reveal that personality factors such as neuroticism, extraversion, openness, agreeableness, and conscientiousness are potential predictors of some rating bias (Kane et al., 1995). In addition, researchers should more closely examine how supervisors' affection for a subordinate, dyadic quality, expectations and history of working together influence the performance appraisal process (Bol & Smith, 2011; Duarte et al., 1994; Hogan, 1987; Judge & Ferris, 1993; Maas et al., 2012). Although several authors have stressed the need to examine gender as an important attribute of social context, only a few field studies have modeled and tested the effects of gender and race composition and supervisor-subordinate similarity on dyadic quality and performance appraisal (Duarte et al., 1994; Elvira & Town, 2001; Varma & Stroh, 2001). Additional work is needed to more fully understand how being a member of a social category produces differential performance evaluation outcomes (Biernat & Sesko, 2013). Minorities are more likely to have supervisors of a different race. So, if differences exist in the performance-evaluation process, ratings may disproportionately disadvantage minority employee outcomes (Elvira & Town, 2001). Until now, little attention has been paid to how a judgment bias can be controlled or reduced as well (Kane et al., 1995; Wong-On-Wing, et al., 2007). The following

debiasing suggestions are left for future research. One might examine whether an education in debiasing strategies during training programs for supervisors reduces judgment biases. Such an education would make the supervisors alert for this potential problem in their evaluations of subordinates and may help to overcome, at least to some extent, their biased judgments (Ghosh & Lusch, 2000; Tan & Jamal, 2001; Tayler, 2010). Also the subjective performance measurement literature in the context of the BSC might be extended. Future research might focus on judgment biases when evaluating based on the BSC: how can employees gain a better appreciation for the measures if they are involved in the selection of those measures and the design of the units' scorecard. This might increase the reliance on all BSC measures, including the unique measures and as such the common measure bias might reduce (Lipe & Salterio, 2000; Libby et al., 2004). Future research could also investigate how different presentation formats and features, such as graphs or traffic lights, facilitate the processing of performance information (Cardinaels & van Veen-Dirks, 2010). Researchers could examine whether the effects of framing the scorecard as a causal chain are stronger when managers are provided with additional causal-chain-related data (e.g., correlations), or when managers are given additional training on using the causal chain (Tayler, 2010). In addition, if supervisors have to justify their evaluation judgments, this accountability could potentially moderate judgment biases (Bol & Smith, 2011). Another fruitful avenue is to look at the consequences of biased performance ratings on subordinate performance (Bol, 2011; Varma & Stroh, 2001). Future empirical research could, for example, examine whether supervisors rate leniently because they expect that leniency bias positively affects subordinate performance or whether they do so to avoid rating costs (Bol, 2011). Studies that focus on common measure bias can explore whether unique non-financial measures are more easily ignored than unique financial measures in a BSC- format. Evaluators surely tend to focus more strongly on financial measures when measures are organized in a BSC-format (Cardinaels & van Veen-Dirks, 2010). If managers tend to ignore non-financial, unique measures in a BSC-format, these performance measures should not be included in the BSC-framework. As subjective performance measures are often non- financial and unique, companies including these subjective performance measures in the BSC may falsely assume managers take those subjective measures into account.

Finally, we discuss future research opportunities related to perceived fairness. Also in this research area there is a need for further research into the role of social, contextual and person-specific factors (Duarte et al., 1994; Dulebohn & Ferris, 1999; Franco-Santos et al., 2012). Researchers should determine which personal and organizational factors impact procedural justice and how procedural justice, in turn, affects organizational outcomes (McFarlin & Sweeney, 1992). Researchers could analyze the effect of different combinations of performance evaluation system design and use on trust and justice perceptions (Hartmann & Slapnicar, 2009, 2012). Does a superior's reputation of trustworthiness have an impact on employee performance in a subjective performance evaluation setting and does this impact differ depending on the specific subjective performance evaluation system design? Or researchers could explore controllability in more detail. Do evaluators respond differently depending on the type of uncontrollable factor (external or internal factor, economic or competitive factor, natural catastrophe, …) (Giraud et al., 2008)? Do subordinates in a subjective performance evaluation setting respond differently to uncontrollable factors if their

superior has a reputation of trustworthiness? Furthermore, future research is needed to define which factors cause favorable or unfavorable supervisor reactions towards influence activities (Wayne & Liden, 1995). Future research could also examine the effects of employees' use of influence tactics on their justice evaluations of the performance evaluation system (Dulebohn & Ferris, 1999).

1.7 Conclusion

In this article we review academic research on subjective performance measurement in high-impact journals. The final selection consists of 67 articles published in 20 high-impact journals during the period 1977 to 2013. We classified this final selection of 67 articles in four research streams: 'optimal contracting', 'discretionary bonus pools', 'judgment biases and debiasing', and 'perceived fairness' and discussed each stream in detail. In the section on optimal contracting benefits and costs related to subjectivity in performance contracts are discussed. Subjectivity can tackle perceived weaknesses in quantitative formulaic bonuses such as incompleteness, short-term focus, susceptibility to manipulation, incentive distortions (congruity issues), risk concerns, environmental uncertainty, uncontrollable factors, moral hazard, asymmetric information or agents "gaming" or manipulating the performance evaluation system. However, subjectivity in performance contracts may cause conflicts and disputes between subordinate and supervisor and judgment biases and evaluation uncertainty may harm employee motivation and productivity. The section on discretionary bonus pools makes clear that this specific type of bonus pools (the magnitude of the bonus pool is based on an explicit formula agreed-upon ex ante and the allocation of the bonus amongst the subordinates is based on supervisor's discretion) can face many of the downsides related to subjectivity in performance contracts, while incorporating the benefits related to subjectivity. The section on judgment biases and debiasing discusses that the introduction of subjectivity in performance measurement introduces different judgment biases by supervisors such as compression bias, centrality bias, biases related to the balanced scorecard, biases related to personal characteristics or biases related to accompanying or competitive information. These judgment biases have an effect on subordinate performance and the perceived fairness of the subjective performance evaluation. The section on perceived fairness shows that perceived fairness is of utmost importance in subjective performance evaluation and that influence activities, favoritism, procedural justice and the (non)-existence of adjustments for uncontrollable factors affect perceived fairness. This article ends with a section on avenues for future research, in which we discuss that more research is needed to uncover the use, benefits and costs related to specific types of subjectivity. Furthermore, we stress that subjectivity should not be studied in isolation, but rather as a part of the whole performance measurement system taking into account the interaction between objective and subjective parts of a compensation contract. Also more research on the role of social, contextual and person-specific factors in the performance evaluation process is needed.

Chapter 2
The role of managerial discretion and manager-employee compensation inequality in manager-employee dyads

Abstract
This paper examines managers' allocation of a bonus to their employees, in terms of their discretion to freely decide on the bonus size and manager-employee compensation inequality in manager-employee dyads. We conduct a 2*2 between-subjects experiment, a multi-period modification of the investment game of Berg et al. (1995) in which we manipulate discretion (low, high) and inequality (low, high). We test and find support for a causal model, in which managerial discretion is negatively associated with the proportion of the bonus pool that a manager allocates to her employee (directly). Furthermore, managerial discretion also indirectly affects the proportion of the bonus pool that a manager allocates to her employee, as it is negatively associated with employee effort and employee effort, in turn, is positively related to the proportion of the bonus pool that a manager allocates to her employee. Additionally, our results suggest that the way in which managers decide is asymmetric due to the presence of manager-employee compensation inequality. Specifically, when compensation inequality is high, managers with high managerial discretion tend to be concerned for fairness, whereas those with low managerial discretion are led by long-term self-interest. This focus on fairness (long-term self-interest) is positively (negatively) associated with the proportion of the bonus pool that a manager allocates to her employee.

Keywords: bonus allocations, managerial discretion, compensation inequality

2.1 Introduction

In this paper, we investigate how a manager (as evaluator and residual claimant) and an employee (being evaluated) experience subjective bonus allocations in manager-employee dyads in environments in which managers have different levels of discretion in determining employee bonus size, and manager-employee inequality varies.[11,12,13] Employee contracts often entail managerial discretion (Choi et al., 2016; Maas et al., 2012; Murphy & Oyer, 2003).

[11] Subjective bonus allocation concerns the common practice in organizations that a manager evaluates and rewards the performance of an employee subjectively (Rajan & Reichelstein, 2006). The correctness of a subjective evaluation cannot be verified by a third party, because it entails subjective judgment (Bol, 2008; Bol & Smith, 2011). Therefore it differs from formula-based bonus allocations.
[12] In this paper we will refer to the manager as female, and to the employee as male in order to improve readability of the paper.
[13] The term 'manager-employee compensation inequality' is a synonym for the term 'vertical pay dispersion' (Guo et al., 2017).

Discretion allows managers to exploit additional, private performance information that is not easily contractible, controllable or foreseeable (Ahn et al., 2010; Bailey et al., 2011; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bol & Smith, 2011; Choi et al., 2016; Gibbs et al., 2004; Rajan & Reichelstein, 2009). It therefore complements objective, formula-based performance evaluation systems (Baker et al., 1994). It reduces dysfunctional incentives, such as employees gaming the evaluation system (Baker et al., 1994; Bol, 2008), and risk for employees (Bol & Smith, 2011; Gibbs et al., 2004; Höppe & Moers, 2011), which consequently leads to better employee incentives (Gibbs et al., 2004; Höppe & Moers, 2011). However, employees' trust in the manager is essential in order to incentivize employee effort (Christ et al., 2012; Gibbs et al., 2004). Managers sometimes act opportunistically by underestimating employee performance in order to pay less reward to the employee or by paying less than the promised reward (Baker et al., 1994; Bol, 2008; Christ et al., 2012; Fisher et al., 2005). This potential manager opportunism causes employees to exert less effort because they fear their effort will not be rewarded (Bol, 2008; Ewing, 2016; Fisher et al., 2005). In order to protect employees against manager opportunism, firms can commit themselves allocating a fixed bonus pool determined by contractible, objective performance measures (Baiman & Rajan, 1995; Bol, 2008; Ewing, 2016; Fisher et al., 2005). Employee effort and employee bonuses are greater if managerial discretion over the *size* of the total employee compensation pool is restricted (Fisher et al., 2005). However, at the same time employee effort and employee bonuses are greater if managerial discretion over the *allocation* of the total employee compensation pool amongst the employees is unrestricted as it reduces employee freeriding (employee opportunism) (Fisher et al., 2005). In this paper we focus on managerial discretion over the *size* of employees' bonus.

Additionally, compensation inequality is an important phenomenon to take into account when investigating subjective performance evaluation. Excessive executive remunerations during the financial crisis, for instance, have provoked a public and political debate on relative pay, social justice and compensation inequality (Bloomberg, 2013, 2014; Reuters, 2013, 2014; The Economist, 2009). In Sweden, The Netherlands and the US, there was a public call for more transparency on the disclosure of compensation inequality e.g. the Dodd-Frank act of 2010 (Bloomberg, 2013; Guo et al., 2017; OECD, 2011b). Guo et al. (2017) show in a budgeting setting that managers are inequity-averse, and therefore more lenient towards employees in a context of manager-employee compensation inequality. It is therefore relevant to investigate how compensation inequality affects subjective bonus allocations and whether managers respond leniently towards their employees in a bonus allocation setting as well.

In this study, we argue that managerial discretion over employee bonus *size* can have beneficial effects in situations with high levels of compensation inequality between the manager and the employee under evaluation, which may amongst others explain why this practice is dominant in organizations. More specifically, we argue that in the presence of compensation inequality, formally limiting managerial discretion reduces the extent to which a manager is concerned for fairness. The reason is that the limited managerial discretion provides the manager with a legitimate excuse to act within the boundaries of these limitations, and therefore makes it more justifiable to behave in a long-term self-interested way, thereby caring less for fairness or inequity-aversion (Tenbrunsel and Messick, 1999). We thereby

contribute to the literature on subjective performance evaluation indicating how social preferences influence managers' use of discretion in bonus allocations (Abernethy et al., 2013; Maas et al., 2012). We also contribute to the literature by further examining the effects of managerial discretion on manager opportunism and employee effort (Fisher et al., 2005).

In our experiment, 140 students interact in pairs of two in a multi-period investment game (Berg et al., 1995). They assume the role of a manager or an employee. The employee exerts effort for the manager by investing experimental units, which determines the size of a bonus pool. Afterwards the manager can divide the bonus pool between the employee and herself. The experiment employs a 2*2 between-subjects design. We manipulate the managerial discretion in determining employee bonus size (*Managerial Discretion*; low, high) and manager-employee compensation inequality (*Compensation Inequality*; low, high). We define *Employee Effort* as the average number of experimental units invested by the employee in the investment game, *Manager Bonus Allocation* as the proportion of the bonus given to the employee and *Manager Fairness* as the extent to which the manager considers a fair bonus allocation. Our experimental results support our hypotheses. *Managerial Discretion* is negatively associated with *Employee Effort* and positively associated with manager opportunism (expressed as the proportion of the bonus kept by the manager). In the high discretion conditions (where employees are not protected against manager opportunistic behavior), employees exert less effort than in the low discretion conditions (under which the employees are protected against manager opportunistic behavior to some extent), which is in line with Fisher et al. (2005). Furthermore, *Employee Effort* is positively associated with the proportion of the bonus given to the employee, as expected along reciprocity theory or tit-for-tat strategy (Axelrod, 1981; Cox, 2004; Falk & Fischbacher, 2006; Fehr et al., 1997; Fehr & Gächter, 2000; Fehr & Schmidt, 1999, 2004; Kreps et al., 1982; Malhotra & Murnighan, 2002; Ockenfels, 2015; Rabin, 1993). Lastly, in case of high manager-employee compensation inequality, managers with high managerial discretion are more concerned about fairness than managers with low managerial discretion. Next, managers that consider fairness allocate a bigger proportion of the bonus to their employee. These finding are important as they show that providing a manager with high managerial discretion over employee bonus size can be worthwhile in case of high fixed manager-employee compensation inequality. In this context, it induces managers to care more about fairness and to consequently reward their employees more generously.

The next section presents the relevant literature and the development of the hypotheses. Section III explains the experimental design and section IV presents the results. In section V we provide the discussion and conclusion.

2.2 Literature & Hypotheses

Compensation Inequality

People frequently have a reasonable good idea about relative compensation inequality between different parties (Hargreaves Heap et al., 2013). Recent public outrage over compensation packages of executives (Bloomberg, 2013, 2014; Reuters, 2013, 2014; The Economist, 2009) and academic research (Anderson et al., 2006; Brülhart & Usunier, 2012;

Greiner et al., 2012; Guo et al., 2017; Hargreaves Heap et al., 2013; Johnson & Mislin, 2011; Smith, 2011) suggest that manager-employee compensation inequality may affect employee effort and the relationship between manager and employee. Excessive executive remunerations have initiated a public and political debate on relative pay and compensation inequality. In a response to the fury of their inhabitants, governments installed regulation and corporate governance policies limiting compensation inequality. In Sweden and the Netherlands the code of good governance dictates that managers should consider the wage gap between themselves and blue-collar workers. Those countries advise to state executive salaries in terms of a multiple of industrial worker salaries (OECD, 2011b). The American counterpart, the Dodd-Frank provision of the Securities and Exchange Commission, forces American public companies to disclose their CEOs' total compensation as a multiple of median total worker pay (Bloomberg, 2013; Guo et al., 2017). Recent initiatives even go beyond simply stating the difference between CEO and average employee compensation. The Swiss could recently vote in a referendum to limit executive pay to maximum 12 times the wage of the least paid employee in the company. Although the referendum was not adopted, legislators clearly think about formally limiting income inequality (BBC, 2013; The Guardian, 2013; The US News, 2013). Given the outlined tendencies, more governmental interventions related to compensation inequality are expected (Daily mail, 2013; OECD, 2011a; Harvard Business Review, 2009).

Academic research addresses compensation inequality as well. According to inequity aversion theory or equity theory, people dislike unfair distribution of wealth relative to the effort each party exerts, the task each party performs or the responsibility each party takes (Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; Festinger, 1954; Hargreaves Heap et al., 2013; Itoh, 2004; Smith, 2011). Rewards should be allocated proportionally to different parties' inputs or work contributions (Adams, 1963, 1965) and people should not perceive any difference between their rewards-to-input ratio and the rewards-to-input ratio of any relevant other (Adams, 1963, 1965; Golman & Bhatia, 2012; Walster et al., 1973).[14] An inequity-averse party feels guilty when his relative payoff is above others' relative payoffs, while that party feels envious when his relative payoff is below that of the others (Fehr & Schmidt, 1999; Itoh, 2004; Johnson & Mislin, 2011). Manager-employee compensation inequality (relative to the task each party performs[15]) may create a social distance (Greiner et al., 2012) and may make an employee envious, which reduces his willingness to exert effort (Dur & Glazer, 2008; Johnson & Mislin, 2011). However, at the same time a manager feels guilty when her payoff is overly large compared to the payoff of her employee and thus with high manager-employee compensation inequality managers are motivated to offer a larger reward to the employee in order to reduce the inequity. Employees might anticipate this inequity averse behavior of the manager and will exert extra effort in order to optimally benefit from the inequity aversion of their manager (Fehr & Schmidt, 1999; Hargreaves Heap et al., 2013; Smith, 2011). Therefore

[14] Equitable reward allocations therefore differ from equal reward allocations, in which the allocated rewards to each employee are the same or similar, irrespective of employees' input or work contributions (Reis & Gruzen, 1976).

[15] In experimental research, participants have been randomly assigned to their roles, they perform similar tasks and they have the same work history or responsibilities. So as these experiments control for the justified reasons for fixed compensation inequality, the presence of a fixed compensation inequality may create a feeling of inequity with the employee (Greiner et al., 2012).

academic findings are mixed: some papers on compensation inequality find a negative effect of compensation inequality on trust or employee effort (Johnson & Mislin, 2011), others find no effect (Anderson et al., 2006; Brülhart & Usunier, 2012; Greiner et al., 2012) and still others find a positive effect (Hargreaves Heap et al., 2013; Smith, 2011). In the typical accounting setting of a budgeting process, Guo et al. (2017) show that managers act indeed more leniently towards employees because of inequity-aversion resulting from manager-employee compensation inequality. Furthermore, employees are more motivated to misreport costs because of compensation inequality.

Managerial Discretion

Firms often allow for managerial discretion in determining employee bonuses (Choi et al., 2016; Maas et al., 2012; Murphy & Oyer, 2003; WorldatWork & Vivient Consulting, 2012). A recent survey indicates that managerial discretion is present in the bonus program of a majority of organizations (WorldatWork & Vivient Consulting, 2012). There are benefits as well as costs related to managerial discretion in the performance evaluation (Bailey et al., 2011; Bol, 2008; Ewing, 2016).[16] Discretion allows managers to use additional, private information that is not easily contractible or quantified in order to adequately incentivize employees (Ahn et al., 2010; Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Gibbs et al., 2004). It allows to incorporate unforeseeable and uncontrollable events that affect employee performance (Bol & Smith, 2011) and it can be used to reward value-enhancing actions that are not easily quantified or measured (Gibbs et al., 2004). With objective performance measures, employees might overly focus on the objectively measured performance dimensions ignoring other, unmeasured performance dimensions or long-term goals (Baker et al., 1994; Gibbs et al., 2004; Ittner et al., 2003). Discretion can protect employees from (downside) risk in their compensation (Bol & Smith, 2011; Gibbs et al., 2004; Höppe & Moers, 2011; Woods, 2012). However, there are downsides to managerial discretion as well. It might introduce favoritism and bias in the performance evaluation (Bol, 2011; Ittner et al., 2003; Lipe & Slaterio, 2000; Moers, 2005; Prendergast & Topel, 1993). Therefore in order to incentivize employee effort it is important that employees trust their managers (Christ et al., 2012; Gibbs et al., 2004), knowing that they might act opportunistically. They can underestimate employee performance in order to pay less reward to the employee or they can be unwilling to pay a promised reward (Baker et al., 1994; Bol, 2008; Christ et al., 2012; Fisher et al., 2005). Employees exert less effort if they fear their effort will not be rewarded (Bol, 2008; Ewing, 2016; Fisher et al., 2005). Managers' reputational concerns can alleviate manager opportunism (Baker et al., 1994; Bol, 2008; Fisher et al., 2005; Kreps et al., 1982; Moers, 2005). Furthermore, firms can limit managerial discretion over the *size* of the bonus pool in order to protect the employee against manager opportunism as well (Fisher et al., 2005). Firms often commit themselves to a fixed bonus pool determined by contractible, objective performance measures (Baiman & Rajan, 1995; Bol, 2008; Ewing, 2016; Fisher et al., 2005; Maas et al., 2012; Moers, 2005). Bonus pools that do not allow for managerial discretion over the *size* of the bonus pool result in more

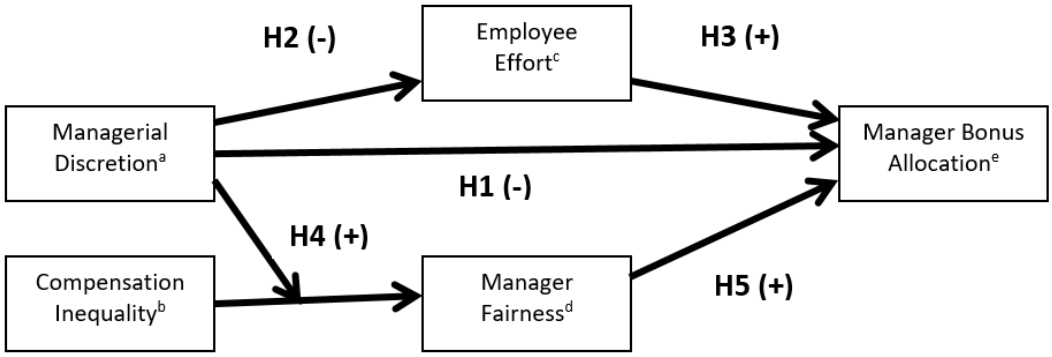
[16] We refer to Bol (2008) and Ewing (2016) for a more extensive discussion of the benefits and costs related to managerial discretion.

employee effort and employee bonuses than bonus pools that have complete discretion over the *size* of the bonus pool (Fisher et al., 2005). However, the *allocation* of the bonus pool across employees is often left at managers' discretion in order to reduce employee opportunism and in order to allow the manager to use her private, non-contractible performance information (Baiman & Rajan, 1995; Bol, 2008; Ewing, 2016; Fisher et al., 2005; Maas et al., 2012; Moers, 2005). Bonus pools that allow managers discretion over the *allocation* of the bonus pool amongst different employees result in more employee effort and employee bonuses (Fisher et al., 2005).

Managerial Discretion and Compensation Inequality: Theoretical Model and Hypotheses

The theoretical model of our study is presented in Figure 2.1. In the context of bonus allocations attached to performance evaluation in manager-employee dyads, this model predicts that the managers' bonus allocation decision is determined jointly by the managers' discretion in determining employee bonus size (hypothesis 1), by employee effort (hypothesis 3) and by managers' concerns for fairness (hypothesis 5).

Figure 2.1 Theoretical Model



[a] *Managerial Discretion* equals 0 in the low managerial discretion conditions, and 1 in the high discretion conditions.
[b] *Compensation Inequality* equals 0 in the low compensation inequality conditions, and 1 in the high compensation inequality conditions.
[c] *Employee Effort* refers to the average number of experimental units of the 100 units endowment the employee invested in the investment project over the 12 periods.
[d] *Manager Fairness* refers to the extent to which a manager is concerned about a fair bonus allocation. Using a seven-point Likert scale, managers rated how important it is for them to be fair.
[e] *Manager Bonus Allocation* refers to the average proportion of the bonus pool that was allocated by the manager to the employee over the number of periods in which there was a bonus pool. If an employee did not invest any experimental units in a particular period, there was no bonus pool and the manager could not allocate any experimental units back to the employee from the bonus pool in that particular period, which happened for 20 manager-employee pairs. We exclude those particular periods as in Smith (2011), because the manager had no choice regarding the amount she offered as a reward to the employee in that particular period.

More specifically, we argue that managers who have limited managerial discretion in determining employee bonus size will offer higher bonus allocations to their employees because the management control system prevents them from acting opportunistically (hypothesis 1). Next, when managers have limited managerial discretion in determining

employee bonus size, these employees will exert more effort because they fear managers' opportunistic behavior to a lesser extent (hypothesis 2). Their managers will reciprocate their higher effort with higher bonus allocations (hypothesis 3). Additionally, a manager who has limited managerial discretion adopts a relatively narrow interpretation of the management control system and is more motivated in her decision-making by long-term self-interest instead of fairness when compensation inequality is high (hypothesis 4), which in turn leads to lower bonus allocations (hypothesis 5). The following paragraphs provide the development of these hypotheses.

Managerial Discretion, Manager Bonus Allocation and Employee Effort

Hypothesis 1 in our model captures the effect of managers' discretion in determining employee bonus size (low vs. high) on the proportion of the bonus pool that a manager allocates to her employee and hypothesis 2 captures the effect of managers' discretion in determining employee bonus size (low vs. high) on employee effort. Managerial discretion falls onto a continuum depending on how vulnerable employees are to the managers' decisions (Ewing, 2016). Even in the absence of formal control systems (such as in the situation of high discretion), managers and employees may cooperate (Abernethy et al., 2013; Baiman & Rajan, 1995; Bolton & Ockenfels, 2000; Bull, 1987; Fisher et al., 2005; Kreps et al., 1982; MacLeod, 2003). Research on investment games shows that individuals are willing to trust other individuals will not follow long-term self-interest, although there are no control systems preventing individuals from acting opportunistically (Anderson et al., 2006; Berg et al. 1995; Brülhart & Usunier, 2012; Ciriolo, 2007; Cox, 2004; Fisher et al., 2005; Greiner et al., 2012; Hargreaves Heap et al., 2013; Johnson & Mislin, 2011; Smith, 2011). However, firms can still choose to install systems that discourage or prevent opportunistic behavior by the manager and that reduce employees' risk of being exploited or betrayed. Strong control systems (Coletti et al., 2005), binding contracts to promote or mandate cooperation (Malhotra & Murnighan, 2002), insurance against betrayal (Lenton & Mosley, 2011), the right to veto managers' bonus allocation (Kanagaretnam et al., 2012) or a limitation of managerial discretion over the size of the bonus pool (Fisher et al., 2005) increase employee effort, trust and cooperation amongst collaborating parties. This experience of control-induced cooperation again has a positive effect on the subsequent level of cooperation, thereby reinforcing the positive effects of the control systems (Coletti et al., 2005).

In conclusion, limiting managerial discretion provides an employee protection against the risk of managers' overly opportunistic behavior i.e. a manager with limited discretion has to provide at least the non-discretionary part of the bonus and therefore cannot act as opportunistically as a manager with high managerial discretion who can freely decide on the bonus allocation. As such, limited managerial discretion leads to more employee effort and higher bonus allocations (Fisher et al., 2005). We therefore formulate our hypotheses as follows:

Hypothesis 1: The proportion of the bonus the manager allocates to her employee is higher when managers' discretion in determining employee bonus size is low than when managers' discretion in determining employee bonus size is high.

Hypothesis 2: Employee effort is higher when managers' discretion in determining employee bonus size is low than when managers' discretion in determining employee bonus size is high.

Employee Effort and Manager Bonus Allocation

Hypothesis 3 in our model captures the effect of effort exerted by the employee on the proportion of the bonus pool that a manager allocates to her employee. According to reciprocity theory, people reward kind actions and punish unkind actions (Cox, 2004; Falk & Fischbacher, 2006; Fehr et al., 1997; Fehr & Gächter, 2000; Fehr & Schmidt, 2004; Malhotra & Murnighan, 2002; McCabe et al., 2003; Rabin, 1993). Reciprocity is thereby considered as a conditional other-regarding preference (Cox, 2004). Generous behavior often induces a reciprocal response e.g. recipients of a gift frequently respond by being generous to those who gave the gift (Fehr et al., 1997; Fehr & Gächter, 2000; Fehr & Schmidt, 2004). Similarly, if an employee exerts effort for a manager, the manager will appreciate this kind behavior and she will likely reciprocate by being more generous in her bonus allocation. This reciprocal response is caused by the cooperative action of the employee as well as the benevolent intentions the manager attributes to this cooperative action (Cox, 2004; Falk & Fischbacher, 2006; McCabe et al., 2003). However, when managerial discretion is limited an employee may cooperate due to the existence of the strong control mechanism instead of due to benevolent intentions (Coletti et al., 2005; Fisher et al., 2005; Lenton & Mosley, 2011; Kanagaretnam et al., 2012). The manager can then attribute the cooperation either to employees' kind intentions (dispositional characteristic) or to the strong control system (situational characteristic). Yet, due to the fundamental attribution error, managers will at least partially attribute the control-induced cooperation to the strong control system (low managerial discretion) instead of to employees' kind intentions, but even in the absence of kind intentions, the control-induced cooperation can elicit mutual cooperation and trust (Coletti et al., 2005). This is in line with the tit-for-tat strategy, which is the winner in Axelrod's prisoners' dilemma tournament (Axelrod, 1980, 1981; Kreps et al., 1982) and which only takes into account people's actions, not their intentions (Rand et al., 2013). "*I'll do to you as you do to me*" is a principle embedded in human nature (Ockenfels, 2015, P. 13). The reasoning above leads us to formulate the following hypothesis:

Hypothesis 3: Employee effort is positively associated with the proportion of the bonus pool the manager allocates to her employee.

Managerial Discretion, Compensation Inequality and Manager Fairness

Hypothesis 4 in our model captures the joint effect of managers' discretion in determining employee bonus size (low vs. high) and the manager-employee compensation inequality (low vs. high) on the extent to which the manager considers fairness in the bonus allocation. Here, we argue that managers with high managerial discretion will respond differently to manager-employee compensation inequality than managers with limited managerial discretion. Fairness is an other-regarding preference that is not conditional on the behavior of the other (Cox, 2004).

When managers care about fairness (or think according to self-centered[17] inequity aversion), they are willing to give up some material payoff to move in the direction of more fair outcomes, as indicated earlier in the paragraph on inequity-aversion theory and equity theory (Adams, 1963, 1965; Bolton & Ockenfels, 2000; Fehr & Schmidt, 1999; Hargreaves Heap et al., 2013; Itoh, 2004; Smith, 2011). When manager-employee compensation inequality is high, a manager with high managerial discretion will be concerned for fairness (Fehr et al., 1997; Fehr & Gächter, 2000; Fehr & Schmidt, 1999, 2004). However, a manager with limited managerial discretion will be concerned for fairness to a much lesser extent. Johnson & Mislin (2011) identified different elements that influence the level of responsibility that a trusted party may feel toward her counterpart. Here we argue that a control system (such as limited discretion) causes a manager to feel less responsible for her employee. The presence of this control system changes the way the manager mentally frames the situation, causing her to perceive the situation as a business decision rather than an ethical decision (Tenbrunsel and Messick, 1999). Control systems may increase cooperation (Coletti et al., 2005) but the basis for this cooperation (i.e., it is economically rational) is different from the basis for cooperation when no control systems are present (i.e. it is the 'right thing' to do) (Tenbrunsel & Messick, 1999). The presence of a control system provides a manager with a legitimate excuse to act within the boundaries of the control system, and therefore induces more focus on long-term self-interest and lowers the concern for fairness. We therefore formulate our hypothesis as follows:

Hypothesis 4: When manager-employee compensation inequality is high, managers with low discretion in determining employee bonus size are less concerned for fairness in the bonus allocation than managers with high discretion in determining employee bonus size.

Manager Fairness and Manager Bonus Allocation

Hypothesis 5 in our model captures the effect of the extent to which the manager is concerned for fairness in the bonus allocation on the proportion of the bonus pool she allocates to her employee. Managers that are concerned for fairness will feel guilty when their relative payoff is above the relative payoff of their employee (Fehr & Schmidt, 1999; Itoh, 2004; Johnson & Mislin, 2011) and therefore these managers will give up material payoff to move in the direction of more equitable outcomes (Fehr & Schmidt, 1999). Managers that are concerned for fairness, will respond by offering a larger bonus to their employee, because they cannot stand inequitable outcomes, whereas managers that are focused on long-term self-interest will keep a larger part of the bonus for themselves (Fehr & Schmidt, 1999). We therefore formulate our hypothesis as follows:

Hypothesis 5: The extent to which managers consider fairness in the bonus allocations is positively associated with the proportion of the bonus pool the manager allocates to her employee.

[17] Inequity aversion is self-centered if people are only interested in the fairness of their own material payoff relative to the payoff of others and do not care per se about the existence of inequity among others (Fehr & Schmidt, 1999).

2.3 Method

Experimental task and design

In order to test our hypotheses, we conducted a 2*2 between subjects experiment.[18] Participants conducted a multi-period variant of the investment game of Berg et al. (1995). We randomly assigned participants to the role of employee (investor or sender in the investment game) and manager (investee or receiver in the investment game) and participants kept their role throughout the experiment.[19] Participants played the variant of the investment game for 12 periods whereby each employee anonymously interacted during each period with the same manager.[20, 21] We manipulated two variables between subjects: *'Manager-Employee Compensation Inequality'* (*Low Inequality or High Inequality*) and *'Managerial Discretion in determining employee bonus size'* (*Low Discretion or High Discretion*).

Experimental Procedures

Each period the employee received an endowment of 100 experimental units. He could decide to invest some (or all) of those experimental units in an investment project or he could decide to keep them. The number of experimental units invested was then tripled. Afterwards the manager decided how many of the tripled experimental units to return to her employee. The employee's payoff then consisted of the experimental units that were not invested and the tripled experimental units that were returned by the manager. The manager's payoff consisted of the tripled experimental units that were not returned to the employee. Additionally, each period the manager and the employee earned experimental units as their fixed compensation. At the end of each period manager and employee got an overview of their own and each other's decisions, payoffs and fixed compensation. The goal for manager and employee was to obtain as many experimental units as possible. Participants were compensated based on the number of experimental units they earned in the experiment. They earned 1.5 euro per thousand experimental units earned in the experiment.

Dependent variables

In this research we focus on managers' and employees' tendency to cooperate. The number of units invested in the investment project represented employees' cooperative behavior. *Employee Effort*$_{it}$ measured how much of the 100 experimental units each employee invested each period. Afterwards, the number of units invested by the employee was tripled representing the return on investment.[22] The manager could then divide the revenues of the investment

[18] We conducted the experiment via the Z-tree experimental software (Fischbacher, 2007).
[19] Here we frame the investment game of Berg et al. (1995) in terms of a labor relationship as in Baker et al. (1994), Schmidt & Schnitzer (1995) and Sloof & Sonnemans (2011).
[20] This is called fixed matching (Hales & Williamson, 2010), which allows for building trust, trustworthiness and reputation. Reputation considerations may mitigate managers' moral hazard problem with respect to paying the bonus (Baiman & Rajan, 1995; Baker et al., 1994; Bol, 2008; Bull, 1987; Fisher et al., 2005; Kreps et al., 1982; Lahno, 1995; Moers, 2005).
[21] The number of periods was unknown to the participants.
[22] The multiplication factor of 3 results in a marginal per capita return (MPCR) of 1.5 (investing 1 unit results in an average profit of 1.5 units for the manager and the employee). The propensity to cooperate increases in the

project between herself and the employee. *Manager Bonus Allocation*$_{it}$ measured for each manager each period the ratio between the number of units the manager returned to the employee and the tripled number of units invested by the employee. This variable represents managers' cooperative behavior. Higher values for this measure indicate lower levels of manager opportunism and vice versa. *Employee Effort (Manager Bonus Allocation*) averages *Employee Effort*$_{it}$ *(Manager Bonus Allocation*$_{it}$*)* over the 12 periods respectively.

Furthermore, in the post-experimental questionnaire we asked participants additional questions. Using a seven-point Likert scale, we asked the managers to indicate how important it is for them to be fair (*Manager Fairness*) and to indicate whether they rewarded the employee because they hoped to earn more money in the long run (*Manager Long-term Self-interest*). *Manager Fairness* refers to the extent to which a manager cares about a fair bonus allocation whereas *Manager Long-term Self-interest* refers to the extent to which a manager is more self-interested when allocating the bonus. Lastly, in the additional analyses we research managers' and employees' trust in each other in greater detail. Using a seven-point Likert scale, employees indicated whether they trusted their manager (*Employee Trust*) and whether they thought their manager trusted them (*Employee Perceived Trust*). Similarly, managers indicated whether they trusted their employee (*Manager Trust*) and whether they thought their employee trusted them (*Manager Perceived Trust*).

Manipulations

In the 2*2 experiment the variables '*manager-employee compensation inequality*' and '*managerial discretion in determining employee bonus size*' were manipulated. We manipulated the variable *manager-employee compensation inequality* at two levels (*Compensation Inequality*; low vs. high) by providing a different fixed compensation to the manager and the employee. Next to the endowment of 100 experimental units that could be invested, the employee received each period a fixed compensation of 150 experimental units, which could not be invested. The manager received each period a fixed compensation of 300 (900) experimental units in the low (high) inequality conditions respectively. As such, we only considered cases in which the employee earned less than his manager, similar to Dur & Glazer (2008) and in line with most real-life employment circumstances.

The variable '*managerial discretion in determining employee bonus size*' was manipulated at two levels (*Managerial Discretion*; low vs. high) as in Bailey et al. (2011). In the high discretion conditions the allocation of the bonus pool was completely left at managers' discretion i.e. the manager decided freely how much of the bonus pool to keep and how much of the bonus pool to give to the employee. In the low discretion conditions the managers could only freely allocate part of the bonus pool (the discretionary bonus part), because a part consisted of a formula-based allocation based on the objective performance measure, '*Employee Effort*' (the non-discretionary bonus part). More specifically, in the low discretion conditions 40% of employees' bonus was non-discretionary and 60% of the bonus was at managers' discretion. The manager with limited discretion decided how much of the bonus

marginal per capita return (Falk & Fischbacher, 2006; Ledyard, 1995). Berg et al. (1995) use the same multiplication factor.

pool to keep and how much of the bonus pool to give to her employee, but her employee got minimum 40% of the bonus pool.[23, 24]

Participants

We recruited 140 undergraduate students from a management accounting course in a large university. These students received a course credit and a monetary reward depending on their performance in the experiment. The average monetary reward equaled 9.40 euro. The experiment lasted between 30 and 60 minutes (depending on the decision speed of the participants). Participants were randomly assigned to a separate computer when they arrived. They received computerized instructions and worked on a task as discussed in greater detail previously. 56% (44%) of the participants was male (female). Participants were 21.5 years old on average (minimum 20 years, maximum 29 years) and they had 8.6 months of work experience, which was significantly different from 0 ($t_{139} = 8.47$; $p < 0.01$).

2.4 Results

Manipulation checks and information sufficiency

We asked a number of questions before the experiment actually started to make sure participants understood the experimental scenario and knew all relevant case information. Nobody could continue with the experiment until all questions were answered correctly. We asked them, for example, which role they played in the experiment, what their fixed income and the fixed income of their counterpart was, whether they could invest any experimental units in the experiment, whether their counterpart could invest any experimental units in the experiment, how many units maximally could be invested in the experiment each period and what the multiplication factor for the investment was. In the low discretion conditions only, we further asked what percentage of the bonus pool the manager should minimally provide as a reward.

In the post-experimental questionnaire we asked additional questions to test our manipulations. In order to test our manipulation of *Managerial Discretion*, participants had to indicate whether the statement "*The manager had to give at least 40% of the return on investment of the investment project back to the employee*" was true or not. All participants in the low discretion conditions answered correctly, while 3 employees in the high discretion conditions incorrectly assumed this statement was true.[25] Additionally we asked employees to

[23] The operationalization of low discretion is similar to the experimental settings of Greiner et al. (2012) and Bailey et al. (2011) and the real-life setting of Abernethy et al. (2013). In Greiner et al. (2012) the manager must allocate at least 90% of the amount invested. In Bailey et al. (2011) 50% of the bonus pool is non-discretionary. In Abernethy et al.'s hospital case study, the managerial discretion refers both to her own bonus as to the bonus of her employee, except that her own bonus is restricted to 3.5 times the average bonus of all her employees.

[24] Rational employees in the low discretion conditions would invest each period 100 experimental units, as this would increase their number of experimental units at least from 100 experimental units to 100 *3 (multiplication factor) * 40% (non-discretionary part of the bonus pool) = 120 experimental units and they might still receive a chunk of the 60% discretionary part of the bonus pool.

[25] We conducted our analyses with and without the participants that missed one or more manipulation checks and found similar results. In this paper we report the results for the data set including all 140 participants, unless specifically mentioned otherwise.

indicate their agreement (on a seven-point scale) with the statement "*The manager could determine the size of my total income to a large extent*". Employees in the high discretion conditions agreed significantly more with the statement than employees in the low discretion conditions ($F_{1, 68} = 5.380$; $p = 0.023$)[26], which is in line with the manipulation of *Managerial Discretion*, as employees fixed income was the same in all conditions. Additionally, the software was programmed such that managers in the low discretion conditions were not able to allocate less than 40% as a bonus.

In order to test our manipulation of *Compensation Inequality*, we asked managers (employees) to indicate whether the statements "*I received as a manager (as an employee) a higher (lower) fixed income than the employee (manager)*" were true or not respectively. All employees answered correctly, while 2 managers answered incorrectly. Furthermore, we asked managers (employees) to indicate on a seven-point scale their agreement with the statements "*My fixed income is strongly different from the fixed income of my employee (manager).*" Managers (employees) in the high inequality conditions agreed significantly more with this statement than managers (employees) in the low inequality conditions ($F_{1, 68} = 34.356$; $p < 0.01$; ($F_{1, 68} = 15.267$; $p < 0.01$)), which is in line with our manipulation of *Compensation Inequality*. Furthermore, employees indicated their agreement (on a seven-point scale) with the statement "*As an employee I had a large influence on the total income of my manager because I decided how many units were invested in the investment project."* Employees in the high inequality conditions did agree significantly less with this statement than employees in the low inequality conditions ($F_{1, 68} = 8.248$; $p < 0.01$). Similarly, managers indicated their agreement (on a seven-point scale) with the statement "*The employee had a large influence on my total income because he decided how many units were invested in the investment project."* Managers in the high inequality conditions did agree marginally significantly less with this statement than employees in the low inequality conditions ($F_{1, 68} = 3.356$; $p = 0.071$). The responses to both statements are in line with the manipulation of *Compensation Inequality*. As the fixed compensation of managers in the high inequality conditions is much higher than in the low inequality conditions, the number of units invested has less influence on the total income of managers in the high inequality conditions than on the income of managers in the low inequality conditions. Thus, our manipulations of discretion and inequality were successful.

Additionally, using a seven-point Likert scale, we asked participants to indicate their agreement with a couple of statements related to the experimental task. Participants indicated they were motivated. The mean response (6.59 on 7) is significantly higher than the scale midpoint of 4 ($t_{139} = 47.83$; $p < 0.01$). The participants said that the instructions were clear. The mean of this measure was 6.18 on 7, which is significantly higher than the scale midpoint of 4 ($t_{139} = 26.20$; $p < 0.01$). Participants indicated that they had enough time to make their decisions (mean of 6.89 on 7, which is significantly higher than the scale midpoint of 4 ($t_{139} = 102.86$; $p < 0.01$)). Together, these results suggest that participants felt they had sufficient time and were capable to perform in the experiment.

[26] Every p-value mentioned in this paper is a two-sided p-value, unless specifically mentioned otherwise.

Descriptive Statistics

Table 2.1 presents the descriptive statistics of our main dependent variables across each of the four experimental conditions. In order to preliminarily test hypothesis 1, we conducted an ANOVA analysis with *Manager Bonus Allocation* as the dependent variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects independent variables. Consistent with hypothesis 1, results (not tabulated) indicate that *Manager Bonus Allocation* is significantly higher when *Managerial Discretion* is low (mean = 56%) than when *Managerial Discretion* is high (mean = 46%) ($F_{1, 66} = 11.772$, $p < 0.01$).[27]

In order to test hypothesis 2, we first conducted a repeated-measures mixed-design ANOVA analysis with *Employee Effort*$_{it}$ as the dependent variable with *Period* (for the periods 1 to 12) as the repeated-measures variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects variables. Consistent with hypothesis 2, results (not tabulated) indicate that *Employee Effort* is significantly higher when *Managerial Discretion* is low (mean = 79.15) than when *Managerial Discretion* is high (mean = 69.52) ($F_{1, 66} = 3.760$; $p = 0.057$).

In order to test hypothesis 4, we conducted an ANOVA analysis with *Manager Fairness* as the dependent variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects variables. Consistent with hypothesis 4, results (not tabulated) indicate a significant interaction of *Managerial Discretion* and *Compensation Inequality* ($F_{1, 66} = 5.836$, $p = 0.018$). Simple effects analysis further indicates a significant effect of *Managerial Discretion* within high inequality ($F_{1, 66} = 6.447$, $p = 0.013$) i.e. under the condition that *Compensation Inequality* is high, *Manager Fairness* is significantly higher when *Managerial Discretion* is high (mean = 6.33 on 7) than when *Managerial Discretion* is low (mean = 5.53 on 7). This analysis also shows a significant effect of inequality within low discretion ($F_{1, 66} = 5.587$, $p = 0.021$). To further investigate our expectations, we will conduct a path analysis.

[27] *Manager Bonus Allocation* refers to the average proportion of the tripled amount invested that was allocated by the manager to the employee over the number of periods in which there was an amount invested. If an employee did not invest any experimental units in a particular period, there was no bonus pool and the manager could not allocate any experimental units back to the employee in that particular period, which happened for 20 manager-employee pairs. We exclude those particular periods as in Smith (2011), because the manager had no choice regarding the amount she offered as a reward to the employee in that particular period. We conducted a repeated-measures mixed-design ANOVA analysis with *Manager Bonus Allocation*$_{it}$ as the dependent variable with *Period* (for the periods 1 to 12) as the repeated-measures variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects variables as well. Due to the 20 missing manager-employee dyads, we are left with 50 dyads. Consistent with hypothesis 1, results (not tabulated) indicate that *Manager Bonus Allocation* is significantly higher when *Managerial Discretion* is low (mean = 56%) than when *Managerial Discretion* is high (mean = 49%) ($F_{1, 46} = 5.622$, $p = 0.022$). Furthermore, the interaction of *Managerial Discretion* and *Compensation Inequality* is marginally significant ($F_{1, 46} = 3.104$, $p = 0.085$).

Table 2.1 Descriptive Statistics: Mean (Standard Deviation)

	Managerial Discretion[a]			
	Low		High	
	Compensation Inequality[b]		Compensation Inequality[b]	
	Low	High	Low	High
Dependent Measure	(n = 18)	(n = 17)	(n = 17)	(n = 18)
Employee Effort[c]	84.07	73.94	70.20	68.88
	(11.77)	(16.39)	(24.98)	(25.27)
Manager Fairness[d]	6.28	5.53	6.00	6.33
	(0.75)	(1.12)	(0.79)	(1.03)
Manager Bonus Allocation[e]	56.26%	56.17%	45.65%	46.33%
	(9.64%)	(10.40%)	(11.33%)	(16.96%)
Employee Trust[f]	4.94	4.53	4.94	4.17
	(1.39)	(1.37)	(1.34)	(1.89)
Employee Perceived Trust[g]	5.00	5.24	5.00	4.61
	(1.14)	(0.90)	(1.00)	(1.46)
Manager Trust[h]	5.28	5.12	5.06	5.33
	(1.13)	(0.99)	(1.64)	(1.08)
Manager Perceived Trust[i]	5.22	5.18	4.88	5.06
	(1.35)	(1.13)	(1.80)	(1.51)

[a] *Managerial Discretion* refers to whether managers' discretion in determining employee bonus size is high (i.e. the bonus allocation is completely left at managers' discretion) or low (i.e. the bonus allocation is partly discretionary and partly non-discretionary).
[b] *Compensation Inequality* refers to whether manager-employee fixed compensation inequality is high or low.
[c] *Employee Effort* refers to the average number of experimental units of the 100 units endowment the employee invested in the investment project over the 12 periods.
[d] *Manager Fairness* refers to the extent to which a manager is concerned about a fair bonus allocation. Using a seven-point Likert scale, managers rated how important it is for them to be fair.
[e] *Manager Bonus Allocation* refers to the average proportion of the bonus pool that was allocated by the manager to the employee over the number of periods in which there was a bonus pool.
[f] *Employee Trust* refers to the extent of trust the employee had in his manager (using a seven-point Likert scale).
[g] *Employee Perceived Trust* refers to the extent of trust the employee estimated his manager had in him (using a seven-point Likert scale).
[h] *Manager Trust* refers to the extent of trust the manager had in her employee (using a seven-point Likert scale).
[i] *Manager Perceived Trust* refers to the extent of trust the manager estimated her employee had in her (using a seven-point Likert scale).

Hypothesis Test

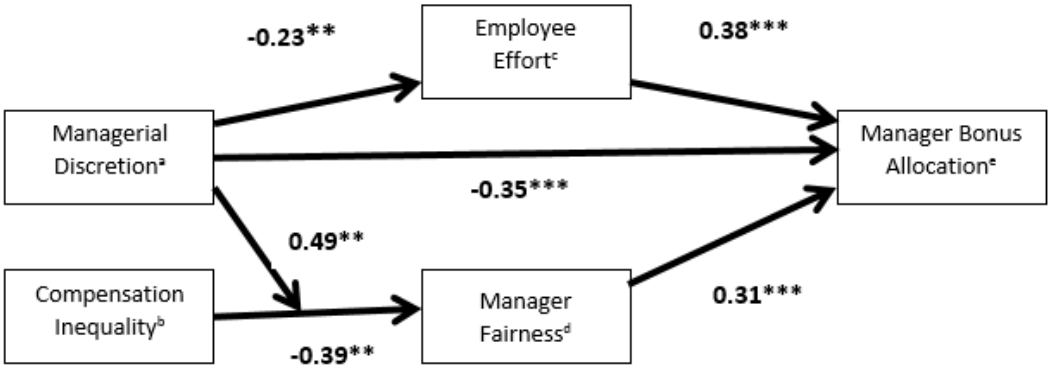
Our theoretical model is depicted in Figure 2.1. We expect a negative relation between *Managerial Discretion* and *Manager Bonus Allocation* (H1) and a negative relation between *Managerial Discretion* and *Employee Effort* (H2). Next, we expect that the interaction between *Managerial Discretion* and *Compensation Inequality* will affect *Manager Fairness*, such that when *Compensation Inequality* is high, managers with low discretion in determining employee bonus size are less concerned for fairness than managers with high discretion in determining employee bonus size (H4). This means we expect a significantly positive interaction. Additionally, we anticipate a positive relation between *Employee Effort* and *Manager Bonus Allocation* (H3) and a positive relation between *Manager Fairness* and *Manager Bonus Allocation* (H5).

To test our model, we employ a path analysis based on structural equation modelling in SAS software. The results (standardized coefficients) of the path analysis are reported in Figure 2.2. The model fits the data well, as indicated by various fit measures ($\chi 2$ (df = 3, n = 70) = 2.6217 (p = 0.4537); Goodness of Fit Index (GFI) = 0.9877; Adjusted GFI (AGFI) = 0.9137; Bentler Comparative Fit Index = 1.0000). Consistent with our expectations, we find support for the hypotheses of our theoretical model.

Additionally, we perform a sensitivity analysis on our causal model by replacing *Manager Fairness* with *Manager Long-term Self-interest* in our path model.[28] Results (not tabulated) for the alternative path are in line with the results in Figure 2.2 and therefore support our reasoning in the theory section. The sign of the path coefficients related to hypotheses 4 and 5 switched from positive to negative, which is logical given we substitute the variable *Manager Fairness* by the variable *Manager Long-term Self-interest*.

[28] As an alternative test for hypothesis 4, we conducted an ANOVA analysis with *Manager Long-term Self-interest* as the dependent variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects variables. Consistent with hypothesis 4, results (not tabulated) indicate a significant interaction of *Managerial Discretion* and *Compensation Inequality* ($F_{1, 66}$ = 6.561, p =0.013). Simple effects analysis further indicates a significant effect of *Managerial Discretion* within high inequality ($F_{1, 66}$ = 6.035, p =0.017) (i.e. when *Compensation Inequality* is high *Manager Long-term Self-interest* is significantly lower when *Managerial Discretion* is high (mean = 4.89 on 7) than when *Managerial Discretion* is low (mean = 6.18 on 7) (H4)) and a significant effect of *Compensation Inequality* within high discretion ($F_{1, 66}$ = 4.494, p =0.038).

Figure 2.2 Path Analysis Results



***, ** and * Indicate p-value, 0.01, 0.05, and 0.1 respectively, two-tailed.
Standardized path coefficients are presented for paths significant at $p < 0.1$, two-tailed.
Fit Indices: χ^2 (df = 3, n = 70) = 2.6217 (p = 0.4537); Goodness of Fit Index (GFI) = 0.9877; Adjusted GFI (AGFI) = 0.9137; Bentler Comparative Fit Index = 1.0000.
[a] *Managerial Discretion* equals 0 in the low managerial discretion conditions, and 1 in the high discretion conditions.
[b] *Compensation Inequality* equals 0 in the low compensation inequality conditions, and 1 in the high compensation inequality conditions.
[c] *Employee Effort* refers to the average number of experimental units of the 100 units endowment the employee invested in the investment project over the 12 periods.
[d] *Manager Fairness* refers to the extent to which a manager is concerned about a fair bonus allocation. Using a seven-point Likert scale, managers rated how important it is for them to be fair.
[e] *Manager Bonus Allocation* refers to the average proportion of the bonus pool that was allocated by the manager to the employee over the number of periods in which there was a bonus pool. If an employee did not invest any experimental units in a particular period, there was no bonus pool and the manager could not allocate any experimental units to the employee from the bonus pool in that particular period, which happened for 20 manager-employee pairs. We exclude those particular periods as in Smith (2011), because the manager had no choice regarding the amount she offered as a reward to the employee in that particular period.

Additional analyses

Next, we also discuss *Manager Bonus Allocation*$_{it}$ (see Table 2.2 and Figure 2.3), *Employee Effort*$_{it}$ (see Table 2.3 and Figure 2.4), the relationships between *Employee Effort*$_{it}$ and *Manager Bonus Allocation*$_{it}$ (see Table 2.4) and *Employee Trust, Employee Perceived Trust, Manager Trust* and *Manager Perceived Trust* in greater detail.

Manager Bonus Allocation$_{it}$,

First, we look at *Manager Bonus Allocation*$_{it}$ in greater detail (see Table 2.2 and Figure 2.3). For each experimental condition, we display the frequency with which managers choose for a particular bonus allocation proportion (or a range of proportions). The observed response pattern can be explained by anchoring[29] on 33.33%, 40%, 50% and 66.67% of the bonus pool.

[29] The anchoring heuristic is a commonly used strategy to divide discretionary bonuses (Bailey et al., 2011). When decision makers use an anchoring heuristic, they adjust their decision insufficiently for additional, decision relevant information, leading to judgments that are biased in the direction of the initial anchor value.

A manager in the high discretion conditions is free to allocate as much of the bonus as she wants. However employees expect to receive at least their initial amount invested, which is 33.33% of the bonus pool (as the initial amount invested is tripled in order to obtain the bonus pool that can be allocated by the manager). Therefore managers' bonus decision in the high discretion conditions gets anchored towards this 33.33% boundary. Of the bonus allocations in the high discretion conditions, 6.43% follows this 33.33% strategy strictly and 14.05% lies in the range [33.33%, 40%[. High discretion managers who provide 33.33% or less as a bonus allocation (< 33.33% & 33.33% in Table 2.2 and Figure 2.3) act opportunistically as they do not allocate any of the profit of the investment to their employee. Those offering even less than 33.33% (< 33.33% in Table 2.2 and Figure 2.3) act extremely opportunistically as the employee even loses money by investing in the investment project. In the high discretion conditions, the manager acts opportunistically in 17.62% of the bonus allocations (< 33.33% & 33.33%) and extremely opportunistically in 11.19% of the bonus allocations(< 33.33%). Managers acting opportunistically are more likely to face no investment in the next period (No investment).

A manager in the low discretion conditions is forced to allocate at least the non-discretionary part of the bonus to her employee, which is 40% of the bonus pool. Therefore managers' bonus decision in the low discretion conditions gets anchored on the non-discretionary part of the bonus (Bailey et al., 2011). Especially, in the low discretion – high inequality condition managers anchor on the 40% rule. 40.68% of the bonus allocations in the low discretion - high inequality condition are in the range [40%, 50%[, which is in line with the theoretical reasoning for hypothesis 4, predicting that managers will use the formal control system as an excuse and will act according to the 40% rule. 9.76% of the bonus allocations in the low discretion conditions follow the 40% rule strictly. Low discretion managers that allocate only 40% of the bonus act as opportunistically as possible and they are more likely to face no investments later on. This opportunistic behavior is located slightly more in the low discretion - high inequality condition (10.29%) than in the low discretion - low inequality condition (9.26%), which results in slightly more cases in which the employee does not invest (No investment) in the low discretion-high inequality condition (2.94%) than in the low discretion-low inequality condition (1.85%).

Furthermore, in all conditions managers are anchoring strictly on an equal split of the bonus pool (50%) (9.05% in the low discretion conditions and 11.67% in the high discretion conditions) or an equal split of the profit taking into account employees' initial investment (66.67%) ((6.19% in the low discretion bonus allocations and 7.62% of the high discretion allocations). Equal division is often considered the social reference point (Bolton & Ockenfels, 2000; Ciriolo, 2007). In the low discretion – low inequality condition, 52.31% of the bonus allocations lie in the range [50%, 66.67%], whereas in the low discretion – high inequality condition only 37.74% of the bonus allocations lie in this range. As indicated before, this is because managers in the low discretion – high inequality condition use the formal control system as an excuse and anchor their bonus allocations to a large extent on 40%. More surprising, in the high discretion – low inequality condition, half of the bonus allocations lies in the range [50%, 66.67%], whereas they could have acted much more opportunistically. Additionally, in the high discretion – high inequality condition 37.49% of the bonus allocations

lies in this range and 7.87% of the bonus allocations is even higher than 66.67%, which is in line with the theoretical reasoning on hypothesis 4.

Lastly, the results on hypothesis 1 already indicated that managers in the low discretion conditions provided a higher proportion of the bonus pool as a reward for their employee than managers in the high discretion conditions (56% versus 46% on average). These findings might not be that surprising at first sight, as the manager needed to return at least 40% in the low discretion conditions, but nonetheless managers in the low discretion conditions allocated on average a 56% bonus, more than the rational 40% boundary ($t_{34} = 9.724$; $p < 0.001$), and even managers in the high discretion conditions offered on average a 46% bonus, well above the 40% boundary as well ($t_{34} = 2.483$; $p = 0.018$). The managers in the high discretion conditions were, of course, not aware of the 40%-rule in the low discretion conditions and 40% as such was no meaningful anchor to them. When managers in the high discretion conditions were able to provide a bonus (i.e. when employees invested units), those managers offered in 73% of the cases a bonus of 40% or higher and only in 27% of the cases a bonus below 40%.[30] So in the majority of the cases high discretion managers already voluntarily followed the 40% rule and only for 27% of the bonus allocations an implementation of a 40% non-discretionary bonus would have made a difference, because it would prevent opportunistic allocation behavior.

[30] Table 2.2 indicates that (1.67% + 18.10% + 11.67% +24.29% +7.62% +4.29%)/(1-7.14%) equals 73% and (11.19% + 6.43% + 7.62%)/(1-7.14%) equals 27%.

Table 2.2 Frequency Table Manager Bonus Allocation

Low Managerial Discretion[a]

Compensation Inequality[b]

Low		High		Total	
Manager Bonus Allocation[c]	Frequency	Manager Bonus Allocation	Frequency	Manager Bonus Allocation	Frequency
No investment[d]	1.85%	No investment	2.94%	No investment	2.38%
40.00%	9.26%	40.00%	10.29%	40.00%	9.76%
]40.00% - 50.00%[	18.98%	]40.00% - 50.00%[	30.39%	]40.00% - 50.00%[	24.52%
50.00%	13.43%	50.00%	4.41%	50.00%	9.05%
]50.00% - 66.67%[	31.94%	]50.00% - 66.67%[	27.94%	]50.00% - 66.67%[	30.00%
66.67%	6.94%	66.67%	5.39%	66.67%	6.19%
> 66.67%	17.59%	> 66.67%	18.63%	> 66.67%	18.10%

High Managerial Discretion

Compensation Inequality

Low		High		Total	
Manager Bonus Allocation	Frequency	Manager Bonus Allocation	Frequency	Manager Bonus Allocation	Frequency
No investment	6.86%	No investment	7.41%	No investment	7.14%
< 33.33%	9.31%	< 33.33%	12.96%	< 33.33%	11.19%
33.33%	8.33%	33.33%	4.63%	33.33%	6.43%
]33.33% - 40.00%[	4.90%	]33.33% - 40.00%[	10.19%	]33.33% - 40.00%[	7.62%
40.00%	0.98%	40.00%	2.31%	40.00%	1.67%
]40.00% - 50.00%[	19.12%	]40.00% - 50.00%[	17.13%	]40.00% - 50.00%[	18.10%
50.00%	13.73%	50.00%	9.72%	50.00%	11.67%
]50.00% - 66.67%[	32.84%	]50.00% - 66.67%[	16.20%	]50.00% - 66.67%[	24.29%
66.67%	3.43%	66.67%	11.57%	66.67%	7.62%
> 66.67%	0.49%	> 66.67%	7.87%	> 66.67%	4.29%

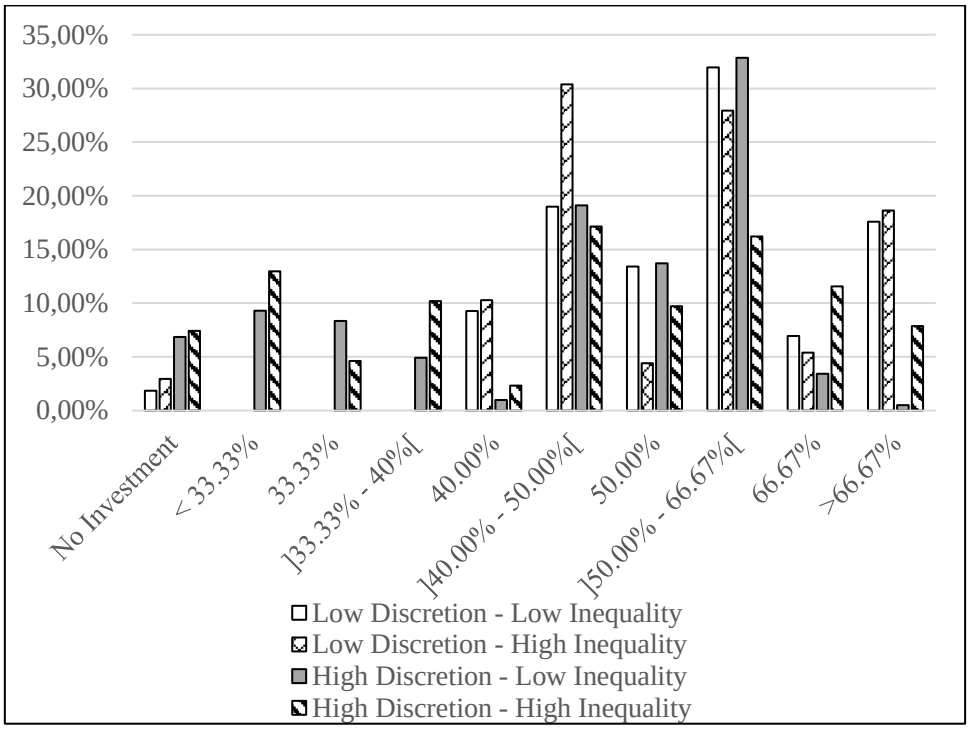
[a] *Managerial Discretion* refers to whether managers' discretion in determining employee bonus size is high (i.e. the bonus allocation is completely left at managers' discretion) or low (i.e. the bonus allocation is partly discretionary and partly non-discretionary).

[b] *Compensation Inequality* refers to whether manager-employee fixed compensation inequality is high or low.

[c] *Manager Bonus Allocation* refers to the proportion of the bonus pool that was allocated by the manager to the employee.

[d] *No investment* represents the cases in which the employee did not invest anything and therefore there was no bonus pool and the manager was not able to allocate any experimental units back to the employee.

Figure 2.3 Frequency Table – Manager Bonus Allocation



Managerial Discretion refers to whether managers' discretion in determining employee bonus size is high (i.e. the bonus allocation is completely left at managers' discretion) or low (i.e. the bonus allocation is partly discretionary and partly non-discretionary).
Compensation Inequality refers to whether manager-employee fixed compensation inequality is high or low.
Manager Bonus Allocation refers to the average proportion of the bonus pool that was allocated by the manager to the employee over the number of periods in which there was a bonus pool.
No investment represents the cases in which the employee did not invest anything and therefore there was no bonus pool and the manager was not able to allocate any experimental units back to the employee.

Employee Effort$_{it}$ and relationship between Employee Effort$_{it}$ and Manager Bonus Allocation$_{it}$,

Next, we look at *Employee Effort$_{it}$* in greater detail (see Table 2.3 & 2.4 and Figure 2.4). In line with hypothesis 2 (managerial discretion leads to less employee effort), employees in the low discretion conditions exert more effort than employees in the high discretion conditions, as employees in the low discretion conditions are protected against overly opportunistic behavior of their managers. Especially in the low discretion-low inequality condition employees invest very large parts of their endowment (59.72% of the investments lie in the range [90, 100]). However, in the other 3 experimental conditions employees invest in 41.67% to 44.29% of the cases a very large part of their endowment ([90, 100]) as well. Furthermore, in the high discretion conditions 28.81% of the investments lie in the range ([0, 50], whereas in the low discretion conditions only 15.48% of the investments lie in this range.

Table 2.3 Frequency Table Employee Effort

Low Managerial Discretion[a]

Compensation Inequality[b]

Low		High		Total	
Employee Effort[c]	Frequency	Employee Effort	Frequency	Employee Effort	Frequency
0	1.85%	0	2.94%	0	2.38%
]0 - 50[	4.63%	]0 - 50[	12.75%	]0 - 50[	8.57%
50	5.09%	50	3.92%	50	4.52%
]50 - 80[	14.81%	]50 - 80[	23.04%	]50 - 80[	18.81%
80	6.94%	80	7.84%	80	7.38%
]80 - 90[	6.94%	]80 - 90[	6.86%	]80 - 90[	6.90%
90	7.41%	90	10.29%	90	8.81%
]90 - 100[	3.24%	]90 - 100[	4.41%	]90 - 100[	3.81%
100	49.07%	100	27.94%	100	38.81%

High Managerial Discretion

Compensation Inequality

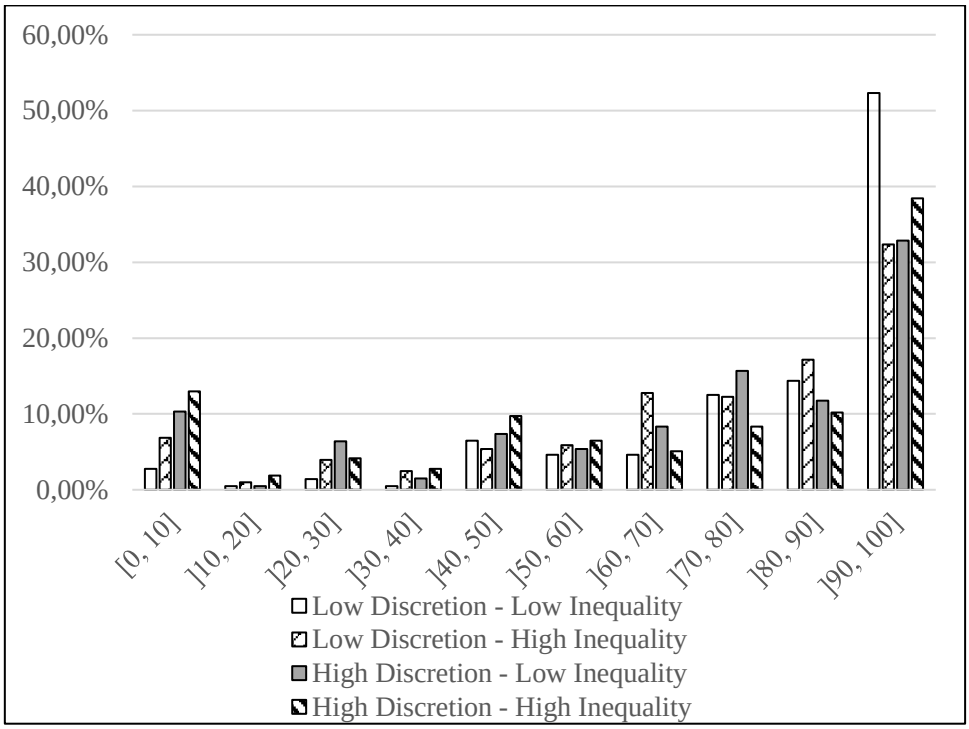
Low		High		Total	
Employee Effort	Frequency	Employee Effort	Frequency	Employee Effort	Frequency
0	6.86%	0	7.41%	0	7.14%
]0 - 50[	12.75%	]0 - 50[	14.81%	]0 - 50[	13.81%
50	6.37%	50	9.26%	50	7.86%
]50 - 80[	22.06%	]50 - 80[	12.50%	]50 - 80[	17.14%
80	7.35%	80	7.41%	80	7.38%
]80 - 90[	2.94%	]80 - 90[	1.85%	]80 - 90[	2.38%
90	8.82%	90	8.33%	90	8.57%
]90 - 100[	2.94%	]90 - 100[	3.70%	]90 - 100[	3.33%
100	29.90%	100	34.72%	100	32.38%

[a] *Managerial Discretion* refers to whether managers' discretion in determining employee bonus size is high (i.e. the bonus allocation is completely left at managers' discretion) or low (i.e. the bonus allocation is partly discretionary and partly non-discretionary).

[b] *Compensation Inequality* refers to whether manager-employee fixed compensation inequality is high or low.

[c] *Employee Effort* refers to the number of experimental units of the 100 units endowment the employee invested in the investment project.

Figure 2.4 Frequency Table – Employee Effort



Managerial Discretion refers to whether managers' discretion in determining employee bonus size is high (i.e. the bonus allocation is completely left at managers' discretion) or low (i.e. the bonus allocation is partly discretionary and partly non-discretionary).
Compensation Inequality refers to whether manager-employee fixed compensation inequality is high or low.
Employee Effort refers to the average number of experimental units of the 100 units endowment the employee invested in the investment project over the 12 periods.

Table 2.4 The effect of managers' opportunistic behavior on Employee Effort

Panel A: High Discretion Conditions

	Model 1			Model 2		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	39.58	4.13	<0.001	73.39	13.84	<0.001
Compensation Inequality	-1.92	-0.27	0.788	-2.05	-0.29	0.770
Period	0.37	0.86	0.392	0.37	0.77	0.439
Manager Bonus Allocation$_{t-1}$	63.59	3.45	0.001			
Manager opportunistic$_{t-1}$				-21.23	-3.56	<0.001
# of observations	384			384		
Overall R^2	0.23			0.25		
Prob > chi2	0.0017			<0.001		

Panel B: Low Discretion Conditions

	Model 3			Model 4		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	42.23	4.60	<0.001	78.67	18.01	<0.001
Compensation Inequality	-9.46	-1.82	0.068	-9.39	-1.90	0.057
Period	1.07	2.80	0.005	1.23	2.98	0.003
Manager Bonus Allocation$_{t-1}$	63.38	4.26	<0.001			
Manager Opportunistic$_{t-1}$				-19.76	-1.98	0.048
# of observations	384			384		
Overall R^2	0.09			0.09		
Prob > chi2	<0.001			<0.001		

We used linear random-effects models and we adjusted the standard errors of the estimates for clustering by manager-employee dyad. 384 observations = (12 periods -1 initial period in which Manager Bonus Allocation$_{t-1}$ was not known yet) * 35 manager-employee dyads – 1 dyad in which the employee did not invest in the first period and Manager Bonus Allocation$_{t-1}$ still was not known in the second period.
The dependent variable is *Employee Effort$_{it}$*, which is number of experimental units the employee invested in period t.
Compensation Inequality equals 0 in the low compensation inequality conditions, and 1 in the high compensation inequality conditions.
Period refers to the period in the investment game.
Manager Bonus Allocation$_{t-1}$ refers to the proportion of the bonus pool that was allocated by the manager to the employee in the most recent period in which the manager could take a decision.
Manager Opportunistic$_{t-1}$ is a dummy variable which is 1 if the Manager Bonus Allocation is 33% or less in the high discretion conditions or 40% in the low discretion conditions in the most recent period in which the manager could take a decision.

In Table 2.4, we conducted multiple linear random-effects models with *Employee Effort$_{it}$* as the dependent variable in order to further investigate the effect of managers' overly

opportunistic behavior on employee effort. In all regressions, we adjusted the standard errors of the estimates for clustering by manager-employee dyad. Panel A focuses on the high discretion conditions, panel B focuses on the low discretion conditions. *Manager Bonus Allocation*$_{it-1}$ is the first proxy for managers' opportunistic behavior. It refers to the proportion of the bonus pool that a manager allocates to her employee instead of keeping the bonus for herself in the most recent period in which the manager could take a decision. Higher values for this variable indicate less opportunistic behavior by the manager. The regression results (Table 2.4, model 1 and 3) clearly indicate that if the manager acts less opportunistically, this leads to more employee effort. If a manager allocates 1% extra to her employee, the employee invests approximately 0.63 experimental units extra in the next period. *Manager Opportunistic*$_{it-1}$ is the second proxy for managers' opportunistic behavior. This is a dummy variable which is 1 if the Manager Bonus Allocation is equal to or less than 33% in the high discretion conditions (Model 2 of Table 2.4) or equal to 40% in the low discretion conditions (Model 4 of Table 2.4) in the most recent period in which the manager could take a decision. The regression results (Table 2.4, model 2 and 4) clearly indicate that if the manager acts opportunistically, this leads to less employee effort. If a manager allocates 33% or less to her employee in the high discretion conditions, the employee invests on average 21.23 experimental units less the next period. If a manager allocates 40% to her employee in the low discretion conditions, the employee invests on average 19.76 experimental units less the next period. This is surprising, as it is rational for employees in the low discretion conditions to exert maximal effort as they would receive a reward that was at least worth 1.2 times their effort (3*40%). Interestingly, though, if the manager acted opportunistically by allocating only 40% of the bonus pool, employees dropped their investments, although this was not the rational thing to do. This is a clear example of negative reciprocity, a punishment strategy in which one party (the employee) incurs a cost to punish another party (the manager) for failing to reciprocate (McCabe et al., 2003). Failures to reciprocate are not forgotten or forgiven and will elicit sanctions (Williamson, 1993).

In the low discretion conditions it is rational to exert maximal effort and for employees in the high discretion conditions it is rational to exert no effort as they risk to lose everything, but it is not per se the optimal strategy. 27 of the 35 employees in the low discretion conditions invested 100 units in one or more experimental periods, but only one employee (in the low discretion-low inequality condition) exerted maximal effort in each period. By doing so, the employee gained only 1960 experimental points. Some of the low discretion employees that did not invest 100 experimental units in the first period or employees that did invest 100 experimental units in the first period, but then occasionally dropped their investments, earned more experimental units than the employee that always invested the maximum amount of 100 experimental units, because their managers provided them with higher bonus allocations in order to persuade them to invest more in the future. For example, 11 of the 35 low discretion employees that did not invest 100 experimental units in each period outperformed the employee that always invested the maximum amount of 100 units. However, low discretion employees were aware of the rational strategy. For example, employees in the low discretion conditions indicated on a seven-point Likert scale their agreement with the statement *"I always invested many units because my supervisor had to give back at least 40% of the bonus pool".* The mean

of this measure was 5.09 on 7, which is significantly higher than the scale midpoint of 4 (t_{34} = 3.56; p <0.01). However, they further indicated on a seven-point Likert scale their agreement with *"If the supervisor gave back only 40% of the bonus pool, I invested the next period less"*. The mean of this measure was 4.77 on 7, which is significantly higher than the scale midpoint of 4 as well (t_{34} = 2.32; p <0.01). This result is in line with the regression results of Table 2.4 and can be explained by negative reciprocity.

Additionally, many employees in the high discretion conditions chose to trust their managers as well. 26 of the 35 employees in the low discretion conditions invested 100 units in one or more experimental periods, but only one employee (in the high discretion-high inequality condition) exerted maximal effort in each period and his manager rewarded this extremely trusting behavior. By doing so this employee gained 2700 points, the best performance of all employees. Next, six of the 35 high discretion employees were able to outperform the low discretion employee that invested 100 experimental units in each period as well.

Employee Trust, Employee Perceived Trust, Manager Trust and *Manager Perceived Trust*

Traditionally, the amount invested in the investment game is considered to reflect employees' trust in the manager, whereas the proportion of the tripled amount invested returned by the manager is considered to reflect managers' trustworthiness towards her employee (Berg et al., 1995).[31] Employees in the low discretion conditions of our experiment did not risk to lose money by investing in the investment game, whereas employees in the high discretion conditions risk to lose money. Therefore, one could argue that employees in the low discretion conditions did not (have to) trust their managers, whereas managers in the high discretion conditions (had to) trust their managers. Managerial discretion in employee performance evaluation falls onto a continuum (Ewing, 2016). By definition, an employee is more vulnerable to managers' decisions (manager opportunism), when managerial discretion increases (Fisher et al., 2005). However, the traditional investment game does not allow one to rule out altruism or inequity aversion as an alternative motive for sending and returning experimental units (Anderson et al., 2006; Cox, 2004). Furthermore, trust is not taking risk per se, but it is rather a willingness to take risk or a willingness to be vulnerable (Mayer et al., 1995). An employee does not need to risk anything in order to be willing to be vulnerable (i.e. to trust a manager). Therefore, an employee under low managerial discretion can trust his manager as much or even more than an employee under high managerial discretion. However, an employee can only show his trusting intentions to others by actually taking risk (instead of being willing to take risk). Therefore, only a high discretion employee displays trusting behavior by investing in the investment game, but this does not mean that low discretion employees did not trust their manager (Mayer et al., 1995). Trust is often confused with cooperation, because trust often leads to cooperation. However, trust is not needed in order to cooperate. People can cooperate with somebody they do not trust (Coletti et al., 2005; Mayer et al., 1995). E.g. strong control mechanisms lead a party to cooperate even in the absence of

[31] Trust is then defined as *"the willingness of a party to be vulnerable to the actions of another party based on the expectation that the other will perform a particular action important to the trustor, irrespective of the ability to monitor or control that other party"* (Mayer et al., 1995).

trust, because they enable the parties to reduce risk and opportunistic behavior (Coletti et al., 2005; Fisher et al., 2005; Kanagaretnam et al., 2012; Lenton & Mosley, 2011). Others can then attribute the cooperation either to the cooperators' innate trustworthiness (dispositional characteristic) or to the strong control system (situational characteristic). However, due to the fundamental attribution error, others at least partially attribute the control-induced cooperation to the strong control system (low managerial discretion) instead of to the cooperators' inherent trustworthiness (Coletti et al., 2005). However, strong control systems can enhance the level of trust among collaborators as well (Coletti et al., 2005; Kanagaretnam et al., 2012; Lenton & Mosley, 2011). Therefore, it is hard to predict in which experimental condition manager-employee dyads trusted each other more. We here investigate employees' and managers' trust in each other in an exploratory way.

In the post-experimental questionnaire we asked employees (using a seven-point Likert scale) whether they trusted their manager (*Employee Trust*) and whether they thought their manager trusted them (*Employee Perceived Trust*). Similarly, we asked managers whether they trusted their employee (*Manager Trust*) and whether they thought their employee trusted them (*Manager Perceived Trust*). T-tests (not tabulated) indicate that each of the four aforementioned variables has a mean that is significantly higher than the scale midpoint of four (at significance level 0.001). This indicates that on average employees and managers in all experimental conditions tended to trust each other. Finally, we conducted four ANOVA analyses with each of these four variables as the dependent variable and *Managerial Discretion* and *Compensation Inequality* as the between subjects variables. We did not observe any differences in trusting beliefs between the different experimental conditions. Managers' and employees' trusting beliefs did not depend on the extent of managerial discretion or compensation inequality.

2.5 Conclusion and discussion

In this paper, we develop and find support for a causal model that explains how managers subjectively allocate a bonus within manager-employee dyads. Specifically, our model explains how managerial discretion in determining employee bonus size and manager-employee compensation inequality influence the proportion of the bonus pool that a manager allocates to her employee (knowing that she can keep the rest for herself). Managers' bonus allocation decision is jointly driven by the management control system (managers' discretion in determining employee bonus size), by other-regarding preferences that are conditional on employees' behavior (reciprocity) and by other-regarding preferences that are not conditional on employees' behavior (fairness). Formal control systems (such as limited managerial discretion) prevent overly opportunistic behavior of managers and therefore lead directly to higher bonus allocations to the employee compared to less formal control systems (that do not limit managerial discretion). Furthermore, formal control systems improve cooperation and employee effort, and in the end managers reciprocate to this extra employee effort by allocating a higher bonus to the employee. However, in case of manager-employee compensation inequality the presence of a formal control system provides a manager with a legitimate excuse to act within the boundaries of the control system. As a result, focusing on long-term self-interest becomes justifiable, in contrast to paying more attention to fairness. This increased

focus on long-term self-interest instead of fairness will ultimately lead to lower bonus allocations to the employee.

These are useful findings for academia and practice, as they provide interesting insights in managers' use of discretion, a common practice in organizations (Bailey et al., 2011; Maas et al., 2012; WorldatWork & Vivient Consulting, 2012). We do not know much about how managers use their discretion to influence employee incentives (Ahn et al., 2010; Bol, 2008). We provide interesting findings, documenting benefits and drawbacks of control systems with explicit incentives to mitigate employer opportunism (Fisher et al., 2005). Especially, they explain why high managerial discretion over employee bonus *size* can be beneficial to employees in situations with high compensation inequality. Earlier research demonstrated the negative effects of managerial discretion over the *size* of employees' bonus on employee effort and employee bonuses (Fisher et al., 2005). Here we indicate that in the presence of high compensation inequality managers with high managerial discretion do not act as opportunistically as earlier thought. In fact, installing a control system that limits managerial discretion in order to reduce managers' opportunistic bonus allocations might backfire in the presence of compensation inequality, as it alters managers' fairness perceptions and causes managers to focus on rationality instead of ethics. In the presence of compensation inequality, managers with high discretion take fairness to a larger extent into account in their bonus decision than managers with low managerial discretion. This insight is novel, since previous academic research suggests that a reduction in managerial discretion in employee bonus *size* is beneficial, as it causes a manager to act less opportunistically, leading to higher employee effort and advantageous employee bonuses (Fisher et al., 2005). However, this literature does not consider manager-employee compensation inequality, a factor that receives substantial attention in the debate on performance rewards today (Anderson et al., 2006; Bloomberg, 2014; Brülhart & Usunier, 2012; Greiner et al., 2012; Guo et al., 2017; Hargreaves Heap et al., 2013; Johnson & Mislin, 2011; Reuters, 2013, 2014; Smith, 2011; The Economist, 2009). Our findings further confirm Fisher et al.'s (2005) results, a reduction in managerial discretion leads to higher employee effort and employee bonuses, but we further add additional insights to the research of Fisher et al. (2005). We contribute to the literature on subjective performance evaluation by indicating how social preferences influence managers' use of discretion in bonus allocations (Abernethy et al., 2013; Maas et al., 2012). Our findings indicate for instance that managers' bonus allocations are both influenced by reciprocity and fairness concerns. Additionally, we demonstrate that managers' mindset towards fairness changes, depending on the extent of managerial discretion in combination with compensation inequality managers balance concerns for long-term self-interest versus fairness in a different way. Next, we contribute to the academic literature on the effect of compensation inequality on fairness, trust or employee effort as well. Some papers on compensation inequality found a negative effect of compensation inequality on trust or employee effort (Johnson & Mislin, 2011), others found no effect (Anderson et al., 2006; Brülhart & Usunier, 2012; Greiner et al., 2012) and still others found a positive effect (Hargreaves Heap et al., 2013; Smith, 2011). In our research we find no effect of compensation inequality on employee effort and trust. However, we answer the call of Anderson et al. (2006) for experimental research that disentangles the effect of fairness, reciprocity and trust in implicit contracts, the call of Ciriolo (2007) and Greiner et al. (2012)

for research on multi-period versions of the investment game and dynamic aspects of trusting relations and the call of Coletti et al. (2005) for research on weaker control systems that provide feedback on collaborating parties actions, without providing strong enough incentives for cooperation (i.e. our high discretion conditions).

As all academic research, this study is subject to a number of limitations, most of them related to the specific formal management control system in place. In the experiment, the number of units invested by the employee was tripled, which led to the bonus pool. In the low discretion conditions 40% of employees' bonus was non-discretionary. In this context it seems rational for employees in the low discretion conditions to exert maximal effort as they would receive a reward that was at least worth 1.2 times their effort (3*40%). Therefore, the findings of H2 (managerial discretion leads to less employee effort) may not be that surprising. More surprisingly though, is that not all employees in the low discretion conditions did exert maximal effort and as such it seems that some ignored certain benefits because of fairness concerns. However, by investing less, these employees were able to convince their managers to provide a larger proportion of the bonus pool to them and by doing so, they outperformed employees that always exerted maximal effort. Employees' reduction in investment was a form of negative reciprocity. The employee incurred a cost in order to punish his manager for overly opportunistic behavior (McCabe et al., 2003). In the low discretion conditions, employees that invested all bonus points would earn 1.2 points per point invested for sure. In reality, bonus pools allocations may not be that benevolent towards their employees, but most incentive systems are designed in such a way that they elicit extra effort and that extra effort is rewarded. The manager needed to return at least 40% in the low discretion conditions, but nonetheless managers in the low discretion conditions allocated on average a 56% bonus, more than the rational 40% boundary, and managers in the high discretion conditions offered on average a 46% bonus. The fact that employee effort acted as a mediator on the relation between managerial discretion and employee bonuses is interesting, as it indicates that managers' bonus allocation was led by conditional other-regarding preferences (reciprocity (H3)), next to the requirements of the management control system (H1). Due to the fact that employees in the low discretion conditions were certain to gain from exerting effort, the underlying motivation for exerting effort of employees in the low discretion conditions was less clear to the managers. Managers in the low discretion conditions could attribute employees' effort to the certain benefits for the employee, to the fact that the employee trusted the manager or to both, while managers in the high discretion conditions could only attribute employees' effort to the fact that the employee trusted the manager. Therefore, managers in the low discretion conditions could have acted in a rational way (allocating only 40% of the bonus pool to the employee) instead of being driven by reciprocity, nonetheless they chose to reciprocate as well. Managers in all experimental conditions indicated that they thought their employees trusted them.

Additionally, companies often allow for managers' justification of bonus allocations which can reduce conflicts, misinterpretations and faulty attributions (Hermans et al., 2017). In addition, feedback can motivate employees as well. These are interesting topics for future research. Similarly, manager and employee had to interact for 12 periods in the experiment. Employees were not allowed to resign if managers acted opportunistically and managers were not allowed to fire lazy employees. Employees' potential threat to leave the company in case

of insufficient bonus allocations and the managers' potential threat of dismissal in case of insufficient employee effort, put a cap on employees' and managers' opportunistic behavior in practice and provide an interesting area for future research. In our paper, 40% of the employee bonus was non-discretionary and this percentage was exogenously decided upon. In practice managers might set this percentage themselves or the percentage might be a result of negotiation between manager and employee. Managers choosing to apply a formal control system (low discretion) or a less formal control system (high discretion) might send a message to their employees, which might lead to different effects than managers that cannot choose the extent of managerial discretion involved. We leave this for future research.

Next, in practice, manager and employee can choose to join companies that fit their preferred work situation regarding manager-employee compensation inequality and managerial discretion and therefore jealousy, envy, inequity-aversion and social preferences may play a lesser role. However, we need future research to confirm these presumptions. Additionally, employee performance and all payoffs in the experiment were very clear to the manager and the employee, while in reality employee performance, compensation inequality, relative payouts of profits and the construction of the bonus system might be more ambiguous and secret. Previous research indicates this affects managers' performance evaluations (Bol et al., 2016) and their use of discretion (Maas et al., 2012). Finally, there was a clear link between bonus amounts received by managers and employees, in that managers and employees had adverse interests in the bonus allocation because the manager was residual claimant, which is not always the case. The manager being the residual claimant in the bonus pool or not, might affect results. We leave this issue for future research as well.

Chapter 3
The role of information accuracy and accountability in bonus allocations

Abstract
Managers often subjectively allocate bonuses amongst their employees in order to stimulate employee performance. Previous literature shows that managers tend to compress their bonus allocations in order to avoid confrontation with their employees, which usually results in disadvantageous organizational outcomes (Bol, 2011, Bol et al., 2016; Moers, 2005). This paper examines the potential role of two control system design elements to decrease managers' compression in bonus allocations: the accuracy of the performance information on which managers base their evaluations and process accountability (meaning that managers get the opportunity to write a justification on how they allocated employee bonuses). A path analysis shows how the two control system elements jointly affect managers' compression in bonus allocations. More accurate performance information leads to more differentiation in bonus allocations, which is in line with the informativeness principle (Holmström, 1979). Furthermore, highly accurate information in the presence of process accountability increases managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations. In this context, managers tend to differentiate more when evaluating employees and allocate bonuses correspondingly.

Keywords: subjective performance evaluation, bonus allocation, information accuracy, process accountability, centrality bias

3.1 Introduction

This study investigates how control system design affects the extent to which managers differentiate in bonus allocations when subjectively assessing the performance of their employees. Subjective performance assessments and attached rewards are highly common in today's work environment in order to create incentives for employees (Rajan & Reichelstein, 2006). A manager can often base her evaluation on a performance overview containing multiple performance measures, such as a balanced scorecard. She then has to subjectively decide upon the weights she attaches to the different performance measures (Cardinaels & van Veen-Dirks, 2010; Höppe & Moers, 2011; Ittner et al., 2003).[32] However, managers'

[32] Therefore subjective performance evaluation differs from formula-based bonus allocations, which are based on weights that are fixed ex-ante. The correctness of a subjective evaluation cannot be verified by a third party, because it entails subjective judgment (Bol, 2008; Bol & Smith, 2011). With subjective evaluations the manager can redirect employees to the performance measures that need more attention in the future due to changing circumstances. Additionally, the manager can avoid that employees manipulate measures or focus too much on particular performance measures ("gaming of the performance evaluation system") (Bol, 2008).

preferences and judgment biases may impair the performance evaluation process (Bol et al., 2016; Moers, 2005; Prendergast & Topel, 1993; Prendergast, 1999), especially when the evaluation decision has direct financial consequences for the employees (Prendergast & Topel, 1993). Managers try to avoid confrontation with dissatisfied employees and they may thereby inadequately differentiate in performance ratings between employees (i.e. centrality bias) (Bailey et al., 2011; Bol, 2011; Bol et al., 2016; Moers, 2005). Centrality bias is related to lower employee performance (Golman & Bhatia, 2012) and lower employee performance improvement (Ahn et al., 2010; Bol, 2011). A fair performance evaluation is desirable for organizations in order to align employee incentives with firm objectives (Baker et al., 1988; Gibbs et al., 2004).[33]

Bol et al. (2016) already demonstrated how the interplay of two elements of the management control system, performance information accuracy and outcome transparency, influences managers' personal costs and benefits related to the performance evaluation and thereby affects managers' compression of employee performance evaluations. More specifically, they show that increasing information accuracy increases the extent to which managers differentiate in the bonus allocation, but only when there is transparency about evaluation outcomes (Bol et al., 2016).[34] In this study, we focus on a complementary element of the management control system, process accountability (meaning that managers get the opportunity to write a justification on how they allocated bonuses to their employees), which is different from outcome transparency (Castilla, 2015).[35] Performance evaluation systems vary across firms, and some firms incorporate the possibility to justify performance assessments, while others do not (Brutus, 2010; Castilla, 2015). While data quality can be different, perceptions of data quality (information accuracy) may vary across its users as well (Pipino et al., 2002; Wang & Strong, 1996). Prior literature already indicates how performance information accuracy and/or process accountability reduce the common measures bias in managers' performance evaluation

[33] Fair performance evaluations are evaluations in which the rewards are allocated amongst the employees proportionally to the employees' input or work contributions. The employees should not perceive any difference between their rewards-to-input ratio and the rewards-to-input ratio of any colleague (Adams, 1963, 1965; Golman & Bhatia, 2012; Walster et al., 1973). Fair performance evaluations therefore differ from equal performance evaluations, in which the allocated rewards to each employee are the same or similar, irrespective of employees' input or work contributions (Reis & Gruzen, 1976). Equal performance evaluations are the result of centrality bias.

[34] The logic behind the result of Bol et al. (2016) is that if outcome transparency is absent, managers will increase the bonus of weaker employees, in line with employees' overly optimistic self-perceptions, in order to avoid costly confrontations with weaker employees. The increased weaker employees' bonuses lead to more compression in the bonus allocation. Being unaware of the bonuses received by their weaker colleagues, the stronger employees will not complain. However, in case of outcome transparency, social comparison between employees is possible and stronger employees will complain about the inflation of weaker employees' bonuses. Managers with relatively accurate performance information will not increase weaker employees' bonuses anymore (which leads to more differentiation in bonuses), as confrontation with weaker employees is less costly in case of accurate performance information, while it allows to satisfy the stronger employees. With less accurate information, managers will still increase weaker employees' bonuses, as this leads to the least confrontation costs.

[35] Supervisors may be accountable for their performance evaluations to their superiors, their subordinates or both (Ferris et al., 2008). However, the justification of performance evaluations to subordinates is the norm (Libby et al., 2004).

(Libby et al., 2004).[36] We investigate the effect of two control system elements, information accuracy and process accountability, because, according to our theory, they jointly affect managers' estimates of employees' perceived procedural fairness of the bonus allocation process and we reason that it is in managers' interest to differentiate more in bonus allocations when they estimate that their employees will perceive the decision-making procedures leading to a differentiation in bonus allocations as more fair.

We first hypothesize that information accuracy is positively related to managers' differentiation in bonus allocation because a manager should put more weight on a measure that is more informative or accurate about employee effort (Holmström, 1979). Additionally, managers take into account that employees evaluate the fairness of the allocation procedures, which then affects employees' perceived fairness of the final bonus allocation (Leventhal, 1980). Employees are more likely to support a detrimental decision when they perceive the decision-making procedures as fair (Bies & Shapiro, 1988). Employees' procedural fairness perceptions of the bonus allocation process will increase if managers base their decision on more accurate information (Colquitt et al., 2001; Leventhal, 1980), if the opinions of employees affected by the decision have been taken into account (Colquitt et al., 2001) and if managers justify their decisions (Bies & Shapiro, 1988; Leventhal, 1980). Similarly, also informational fairness perceptions are related to information accuracy and the possibility for justification (Colquitt, 2001).[37] More specifically, accountability leads managers to engage in extra information processing in search for reasons to justify their decisions. Consequently they make better and more fair decisions, taking into account the perceptions of the employees affected by the decisions (Ashton, 1990; Brutus, 2010; Gibbins & Newton, 1994; Kennedy, 1993; Lerner & Tetlock, 1999; Libby et al. 2004; Mero & Motowidlo, 1995; Mero et al., 2003). Providing a justification allows managers to signal they act fairly and it may further increase employees' perceived procedural fairness of the bonus allocation process (Bies & Shapiro, 1988). However, the effect of accountability on employees' perceived procedural fairness is likely to depend on performance information accuracy. If the data quality of the performance measures is (perceived as) sufficiently high to make a decent decision and to write out a justification, employees receiving a justification of the bonus allocation process will have higher procedural and informational fairness perceptions (Colquitt et al., 2001; Leventhal, 1980; Libby et al., 2004; Pipino et al., 2002; Wang & Strong, 1996; Yim, 2001). In contrast, when the accuracy of the performance measures is (perceived as) low, it is less possible to make a decent decision and come up with a clear justification. In this case, employees procedural and informational fairness perceptions will be lower. Our second hypothesis therefore states that when managers have the opportunity to write out a justification, they have higher estimates of employees' procedural fairness perceptions and employee acceptance of differentiation in bonus allocations when information accuracy is high than when information accuracy is low. Additionally, when employees have higher perceived procedural fairness, they

[36] The common measures bias refers to managers' tendency to focus on performance measures common to different business units and to ignore performance measures unique to a particular business unit while evaluating performance (Libby et al., 2004; Lipe & Salterio, 2000).

[37] Informational fairness perceptions refer the extent to which employees perceive that they receive timely, accurate, and reasonable explanations about decision-making processes or outcomes (Colquitt, 2001).

will more easily accept the final bonus allocations (Bies & Shapiro, 1988; Leventhal, 1980). This opens the path to the manager to differentiate more in her bonus allocations, as managers try to rate employees' performance in line with their inputs and work contributions (Adams, 1963, 1965; Fehr & Schmidt, 1999; Golman & Bhatia, 2012; Maas et al., 2012; Walster et al., 1973). Thus, our third hypothesis predicts that managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations are positively related to the managers' differentiation in bonus allocation.

We use a 2*2 between-subjects experiment manipulating the accuracy of the performance measures (high vs. low) and the possibility managers get to provide a written justification for their bonus allocation (absent vs. present). In our experiment, students participate as evaluators in a hypothetical case setting. They assume the role of a regional manager allocating a fixed bonus pool amongst five store managers based on a performance overview containing three performance measures. Our findings, based on a path analysis, indicate that increasing performance information accuracy increases the managers' differentiation in bonus allocations in a direct way (hypothesis 1). Next, the path analysis shows that information accuracy interacts with accountability such that accountability increases managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations, but only when information accuracy is relatively high (hypothesis 2). Finally, managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations are positively related to the managers' differentiation in bonus allocations (hypothesis 3).

This study contributes to the accounting and management literature on compressed rating behavior in bonus pools (Ahn et al., 2010; Bailey et al., 2011; Bol, 2011; Bol et al., 2016; Golman & Bhatia, 2011; Levin, 2003; MacLeod, 2003; Moers, 2005; Prendergast & Topel, 1993) in several ways. First, we contribute to the literature indicating that choices related to the design of the management control system should not be made in isolation, as they jointly influence evaluation bias (Bol et al., 2016; Libby et al., 2004). We research a similar setting as in Bol et al. (2016) and our results are strongly in line with theirs. However, we research a different kind of transparency (Castilla, 2008, 2015). Bol et al. (2016) focus on outcome transparency, a forced disclosure of the outcomes of the bonus allocation to all employees, which likely affects employees' perceived distributive fairness and potential for social comparison. This paper researches process accountability, a voluntary disclosure of the bonus allocation process to each individual employee separately, but not the disclosure of outcomes of the bonus allocation. Additionally, we provide more details about the underlying process leading to compressed rating behavior, by examining how the two control system elements, information accuracy and process accountability, jointly affect managers' estimates of employees' procedural fairness perceptions and consequently managers' performance evaluation.

Second, we contribute to this literature by investigating the allocation of a fixed bonus pool. Prior literature focuses on the allocation of independent bonuses and shows that managers tend to rate asymmetrically: they often inflate bonuses of weaker employees in order to avoid confrontation with them, which results in compression in the bonus allocation (Bol, 2011; Bol et al., 2016; Moers, 2005). However, given the zero-sum nature of fixed bonus pools, any bonus

decision to certain employees involves some kind of balancing decision towards the other employees (Baiman & Rajan, 1995; Bol et al., 2016; Rajan & Reichelstein, 2006). We show that managers' decision to compress bonus allocations with a fixed bonus pool is symmetric in the sense that managers tend to inflate the bonuses of weaker employees as well as to decrease the bonuses of stronger employees. Lastly, we add to the literature on accountability in performance evaluations (Kennedy, 1993; Mero et al., 2003; Mero et al., 2007; Shore & Tashchian, 2002). In this paper we show that the effect of accountability differs depending on the level of accuracy of the performance measurement system. We hereby provide further insights into the use and effect of written comments in subjective performance assessments (Brutus, 2010).

The next section presents the relevant literature and the development of the hypotheses. Section III discusses the research method and section IV focuses on the results. In section V we provide the discussion and conclusion.

3.2 Literature & Hypotheses

Subjective performance evaluations and differentiation in bonus allocations

Subjective performance evaluations of employees by their manager are omnipresent (Gibbs et al., 2004; Golman & Bhatia, 2012; Rajan & Reichelstein, 2006). For the evaluation, a manager typically uses a performance overview containing multiple performance measures, often presented as a balanced scorecard. The operationalization of the performance scorecards varies considerably across firms (Cardinaels & van Veen-Dirks, 2010; Lohman et al., 2004), as the performance measures in the scorecard are dedicated to key aspects of the business, reflecting the company's' strategy (Banker et al., 2004; Ittner et al., 2003, Libby et al., 2004). The manager then subjectively weighs performance on the different performance measures (Höppe & Moers, 2011; Ittner et al., 2003), but there is no best way to determine these weights (Cardinaels & van Veen-Dirks, 2010; Libby et al., 2004). Managers might ignore certain performance measures, change the weightings of performance measures from period to period or include factors that were not even performance measures (Ittner et al., 2003; Lipe & Salterio, 2000, 2002). Furthermore, evaluators might overly weigh performance on particular categories of measures at the expense of other categories, e.g. performance measures common to different business units versus measures unique to a particular business unit (Banker et al., 2004; Libby et al., 2004; Lipe & Salterio, 2000), financial measures versus non-financial measures (Cardinaels & van Veen-Dirks, 2010; Ittner et al., 2003) or strategically linked performance measures versus non-linked measures (Banker et al., 2004; Humphreys & Trotman, 2011). Additionally, managers often do not adequately weigh differences in performance measures, resulting in insufficient differentiation in employees' performance ratings i.e. centrality bias (Bailey et al., 2011; Bol, 2011; Bol et al., 2016; Golman & Bhatia, 2012; Levin, 2003; MacLeod, 2003; Moers, 2005). Managers try to avoid disputes and conflicts with their employees and they make a tradeoff between providing incentives to the employee ex ante (MacLeod, 2003) and reducing the cost of conflict with the employee ex post (Bol, 2011; Bol

et al., 2016; MacLeod, 2003).[38] Fair, unbiased performance evaluations are needed in order to align employee incentives with firm objectives (Gibbs et al., 2004). Biased performance evaluations such as centrality bias reduce employees' incentives (Ahn et al., 2010; Baker et al., 1988; Bol, 2011; Golman & Bhatia, 2012), which leads to lower employee productivity (Baker et al., 1988), lower employee performance (Golman & Bhatia, 2012) and lower employee performance improvement (Ahn et al., 2010; Bol, 2011). Biased performance evaluations make it more difficult to distinguish good performers from average or bad performers as well, and therefore they may cause promotion of the wrong employees or misidentification of employees' training needs (Bol, 2008).

The performance evaluation's success depends crucially on the confidence of the employees in the evaluator (Gibbs et al., 2004) and their attitudes related to the fairness and acceptability of the performance evaluation system (Dickinson, 1993, p. 142; Landy et al., 1978; Lawler, 1967). Earlier research already examined the effect of contextual factors and components of the performance evaluation system on employees' reactions related to fairness or justice of the systems (Dickinson, 1993; Dulebohn & Ferris, 1999; Hartmann & Slapnicar, 2012; Landy et al., 1978), such as the amount of the employees' voice in the performance evaluation process (Hartmann & Slapnicar, 2012).

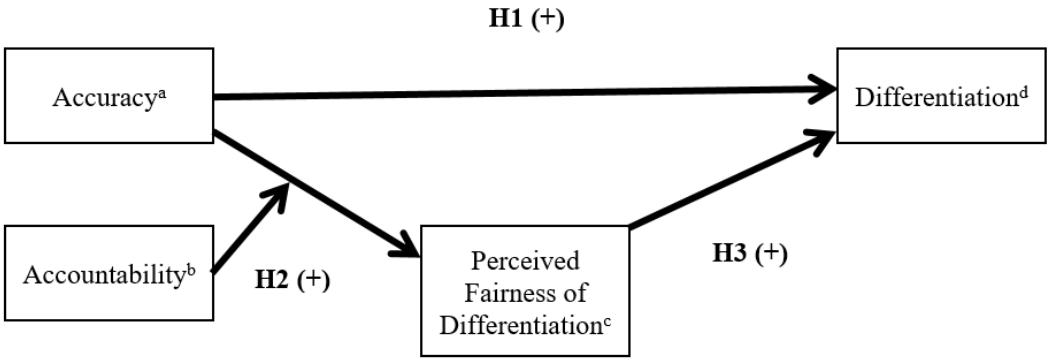
In this study, we examine how variations in the perceptions of the accuracy of the performance information on which the manager bases her evaluation and process accountability affect the manager's estimates of employees' procedural fairness perceptions of differentiation in bonus allocations and how these estimated fairness perceptions consequently affect compression in bonus allocations.

Information Accuracy and Process Accountability: Theoretical Model and Hypotheses

The theoretical model of our study is presented in Figure 3.1. This model predicts that the interaction of performance information accuracy and process accountability determines managers' estimates of employees' procedural fairness perceptions (hypothesis 2). Managers' differentiation in the bonus allocation is then determined jointly by performance information accuracy (hypothesis 1) and their estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations (hypothesis 3).

[38] Traditionally, managers inflate performance ratings of weak employees in an attempt to avoid confrontation with unsatisfied, poor performers, while maintaining good relationships with stronger performers (Bol, 2011; Bol et al., 2016; Colella et al., 2007; Golman & Bhatia, 2012; Mero et al., 2007; Shore & Tashchian, 2002). This asymmetric inflation of the ratings results in less variance in the remaining performance ratings i.e. performance evaluation compression (Bol et al., 2016). However, in the context of a fixed bonus pool, as in our study, these arguments do not hold, as other employees are affected by the bonuses given to colleagues. Increasing the bonus of one employee leads automatically in a decrease in the bonuses allocated to one or more other employees. With a fixed bonus pool, allocating bonuses fairly, in line with individuals' performance or work contributions (Adams, 1963; 1965), is the best strategy for a manager in order to avoid confrontation with employees (Walster et al., 1973).

Figure 3.1 Theoretical Model



[a] *Accuracy* refers to whether the accuracy of the performance information is high or low. Accuracy is coded as 1 for the High Accuracy conditions and 0 for the Low Accuracy conditions.
[b] *Accountability* refers to whether the manager had an opportunity to provide a written justification for the bonus allocation or not. Accountability is coded as 1 for the Accountability conditions and as 0 for the No Accountability conditions.
[c] *Perceived Fairness of Differentiation* reflects the *managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations*. Using a seven-point Likert scale, managers rated their agreement with the following statement: *"If I would offer a well-performing store manager a much higher bonus than a badly-performing store manager, then my store managers would think that I am fair".*
[d] *Differentiation* refers to the extent of differentiation in the bonus allocation. It is measured by two alternative variables (i.e. *Differentiation*, the standard deviation of the bonuses given to each of the 5 store managers or *Bonus Range*, the maximum bonus allocated to a store manager minus the minimum bonus allocated to a store manager).

More specifically, we argue that managers receiving more accurate performance information will differentiate more in their bonus allocations, in line with Holmström's (1979) informativeness principle (hypothesis 1). Furthermore, if the performance measures are sufficiently accurate to make a motivated decision and write out a decent justification, managers estimate that employees receiving a justification of the bonus allocation process will have higher procedural and informational fairness perceptions (Colquitt et al., 2001; Leventhal, 1980; Libby et al., 2004; Pipino et al., 2002; Wang & Strong, 1996; Yim, 2001). In contrast, when the accuracy of the performance measures is (perceived as) low, it is less possible to make a decent decision and write out a corresponding justification. Consequently, managers' estimates of employees procedural and informational fairness perceptions will be lower (hypothesis 2). Additionally, higher perceived procedural fairness results in a more easy acceptance of final bonus allocations (Bies & Shapiro, 1988; Leventhal, 1980). This allows the manager to differentiate more in bonus allocations, in line with her fairness preferences (Adams, 1963, 1965; Fehr & Schmidt, 1999; Golman & Bhatia, 2012; Maas et al., 2012; Walster et al., 1973) (hypothesis 3). We discuss each of the three hypotheses in greater detail below.

Information accuracy and Differentiation in bonus allocations

An increase in the perceived quality of the performance information (information accuracy) might be a useful means to reduce centrality bias. When evaluating employees, a manager subjectively weighs performance on the different performance measures (Höppe & Moers,

2011; Ittner et al., 2003). The determination of exact weights is often impossible (Libby et al., 2004). However, the informativeness principle proclaims that any measure that is informative of employee effort should be incorporated in the performance evaluation and a manager should put more weight on a measure when that measure is more informative or accurate (Holmström, 1979).[39] Zero weights are unsuitable (Ittner et al., 2003; Libby et al., 2004) unless the data quality is sufficiently poor (Libby et al., 2004; Yim, 2001). Centrality bias is (amongst others) caused by the uncertainty about employees performance (Golman & Bhatia, 2012). Managers try to rate employees' performance fairly, in line with employees' inputs and work contributions (Adams, 1963, 1965; Golman & Bhatia, 2012; Maas et al., 2012; Walster et al., 1973). However, when the accuracy of the performance measures is (perceived as) low, it is difficult to judge how well an employee performed and to determine a fair rating. Managers want to decrease the probability that the subjective performance evaluation is very different from the true employee performance level. Taking into account that the probability for extremely good or bad employee performance is low, managers try to avoid these extremes, which then means that they do differentiate to a lesser extent in the performance evaluation (Bol, 2011; Golman & Bhatia, 2012; Levin, 2003; MacLeod, 2003). In contrast, when the accuracy of the performance measures is (perceived as) high, it is easier to judge how well an employee performed and to come to a fair, differentiated rating. We therefore formulate the following hypothesis:

Hypothesis 1: Information accuracy is positively related to managers' differentiation in bonus allocation.

Accountability, Information accuracy and Managers' estimates of employees' procedural fairness perceptions

Next, we research the possibility that managers' estimates of employees' procedural fairness perceptions are jointly influenced by data quality (information accuracy) and process accountability. In order to make adequate performance assessments, managers need to invest time and effort in gathering and processing employee performance information. Since managers' time is scarce and costly (Bol, 2008, 2011; Bol et al., 2016; Kramer & Maas, 2014; Maas et al., 2012), they might not dedicate enough time to the information processing (Bernardin & Villanova, 1986; Bol, 2011). Prior literature indicates that the requirement to justify one's decision to others (accountability) leads to better decision making and lower information processing biases such as overconfidence or insensitivity to new information (Ashton, 1990; Kennedy, 1993; Libby et al. 2004). Accountable managers are more attentive, engaged and motivated in the evaluation process. They care about how others see them (Mero et al., 2006) and they want to avoid appearing incompetent in front of the persons to whom they are accountable. Therefore they engage in extra effort and self-critique in search for reasons to justify their decisions (Lerner & Tetlock, 1999; Tetlock et al. 1989). Accountable managers will look for cues about preferences of the persons to whom they are accountable

[39] Information accuracy refers to the extent to which information is informative about employees' effort. It refers to the variability around a point estimate of an employee's effort level (Bol et al., 2016). Information accuracy is similar to information precision (Banker & Datar, 1989).

and, when possible, provide ratings that are consistent with the perceived views (Mero et al., 2007) in order to maximize the defensibility of their decision (Frink & Ferris, 1998; Klimoski & Inks, 1990; Mero & Motowidlo, 1995; Mero et al., 2007; Tetlock et al., 1989). Consequently they make better and more fair decisions, taking into account the opinions of the employees affected by the decisions (Ashton, 1990; Brutus, 2010; Gibbins & Newton, 1994; Kennedy, 1993; Lerner & Tetlock, 1999; Libby et al. 2004; Mero & Motowidlo, 1995; Mero et al., 2003).

Employees' procedural fairness perceptions of the bonus allocation process are positively affected when the opinions of various employees affected by the evaluation have been taken into account (Colquitt et al., 2001) and when managers justify their decisions (Bies & Shapiro, 1988; Leventhal, 1980). Providing a justification allows managers to signal they act fairly and it may further increase employees' perceived procedural fairness of the bonus allocation process because it eliminates a worst-case reading of the managers' motives and intentions (Bies & Shapiro, 1988). Furthermore, informational fairness perceptions are positively influenced by the possibility for justification as well. However, we expect that this is only the case if the performance measures are perceived as sufficiently accurate to make a proper decision (Golman & Bhatia, 2012; Libby et al., 2004; Pipino et al., 2002; Wang & Strong, 1996; Yim, 2001). It means that the positive effect of accountability on employees' perceived procedural fairness is likely to depend on the accuracy of the information about employees' performance used in the decision-making (Colquitt et al., 2001; Leventhal, 1980). Managers try to rate employees' performance according to their fairness preferences (Adams, 1963, 1965; Golman & Bhatia, 2012; Maas et al., 2012; Walster et al., 1973). However, in case of inaccurate performance information, it is difficult to judge how well an employee performed and to determine a fair rating. In this situation it is hard for a manager to provide a reasonable justification of the performance evaluation and to credibly signal fair rating behavior to an employee. In this case, employees' procedural and informational fairness perceptions will be lower. For managers with accurate performance information, on the other hand, it is easier to reasonably explain the performance evaluations and to credibly signal fair rating behavior to the employees. As a consequence, employees receiving a justification for the bonus allocation process will have higher procedural and informational fairness perceptions (Colquitt et al., 2001; Leventhal, 1980; Libby et al., 2004; Pipino et al., 2002; Wang & Strong, 1996; Yim, 2001). This reasoning leads to the following hypothesis:

Hypothesis 2: Accountability will increase managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations, but only when information accuracy is relatively high.

Managers' estimates of employees' procedural fairness perceptions and Differentiation in bonus allocations

The performance evaluation's success depends crucially on employees' attitudes related to the fairness and acceptability of the performance evaluation system (Dickinson, 1993, p. 142; Landy et al., 1978; Lawler, 1967). Employees are more likely to support a detrimental decision when they perceive the decision-making procedures as fair (Bies & Shapiro, 1988). Managers take into account that employees evaluate the fairness of the allocation procedures and that these evaluations affect the perceived fairness of the final bonus allocation (Leventhal, 1980).

We therefore reason that it is in managers' interest to differentiate more in bonus allocations when they estimate that their employees will perceive the decision-making procedures leading to a differentiation in bonus allocations as more fair. This results in the following hypothesis:

Hypothesis 3: Managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations are positively related to the managers' differentiation in bonus allocations.

3.3 Method

Experimental task and design

In order to test our hypotheses, we conducted a 2*2 between subjects experiment. Participants conducted a performance evaluation task in a hypothetical case setting. They assumed the role of a regional manager of a retail chain. This regional manager had to divide a fixed bonus pool of 50.000 euros entirely amongst five store managers based on three performance measures. We manipulated two variables between subjects: the accuracy of the performance measures (*High Accuracy or Low Accuracy*) and whether the participants had the opportunity to justify or clarify their performance evaluations towards each of the five store managers (*Accountability or No Accountability*).

Experimental procedures

Upon arrival at the computer room each participant was randomly assigned to one of the four experimental conditions. Participants received computerized instructions on the experimental task.[40] They assumed the role of a regional manager of a local retail chain that was responsible for the five stores in her region. The instructions specified that these five stores had a similar size, sold the same assortment and faced similar market circumstances. Furthermore, these stores did not compete directly with each other. Participants were informed that the retail chain recently installed a bonus system whereby store managers of well performing stores could earn an extra financial reward. The instructions noted that the aim of the bonus system was to encourage store managers to perform as well as possible. The regional manager got the duty to allocate the bonus of 50.000 euros amongst the five store managers. The performance evaluation had to be based upon each store's scores on three performance measures: employee satisfaction, customer satisfaction and mystery shopping. Each component was scored on a 0-100 scale. The instructions specified that a survey amongst employees (customers) measured employee (customer) satisfaction. Additionally, it was stated that head office sent a mystery shopper to each store. The mystery shopper valued store performance in a subjective score. The case stipulated that the performance on the three criteria (employee satisfaction, customer satisfaction and mystery shopping) would provide an indication of overall performance for each store manager in the previous year and that participants had to base their bonus decision on each of the three performance criteria. Participants in all four conditions received the same performance overview for the five store

[40] We conduct the experiment in the Ztree experimental software (Fischbacher, 2007).

managers containing the same performance information on the three performance measures.[41] Table 3.1 displays the performance overview.

Table 3.1 Performance overview for 5 store managers and 3 performance measures

	Employee satisfaction (score from 0 to 100)	**Customer satisfaction (score from 0 to 100)**	**Mystery shopping (score from 0 to 100)**
Store manager A	68	72	79
Store manager B	73	71	71
Store manager C	95	85	92
Store manager D	78	69	64
Store manager E	51	48	49
Average score	**73**	**69**	**71**

The table displays the performance overview that all participants received during the experiment. Participants were responsible for 5 store managers (Store manager A to E). They had to allocate a bonus amongst their store managers based on the performance on the 3 performance measures that were equally important and that were scored on a scale from 0 to 100 (employee satisfaction, customer satisfaction and mystery shopping). Store manager C performed the best on all three performance measures, store manager E performed the worst on all three performance measures and the store managers A, B and D performed somewhat in-between.

Participants were informed that the rewards allocated to each store manager would be personal and confidential. After reading all case information, participants were asked to divide the bonus of 50.000 euros amongst the five store managers. They then received a number of post-experimental and demographic questions.

Dependent variables

In this research we focus on the manager's tendency to insufficiently differentiate bonus allocations (centrality bias). We use two measures to capture the extent to which the regional manager differentiates in the bonus allocation to the store managers: *Differentiation* and *Bonus Range*. We measured the bonus amounts (of the 50.000 euros fixed bonus pool) participants allocated to each of the five store managers. *Differentiation* is defined as the standard deviation of the bonuses given to each of the five store managers.[42] *Bonus Range* is defined as the maximum bonus allocated to a store manager minus the minimum bonus allocated to a store

[41] Firms often use scorecards that only contain performance measures that are common to all business units (Cardinaels & van Veen-Dirks, 2010).

[42] This measure is similar to the measure for centrality bias in Bol (2011). In her paper the ratio between the standard deviation of the objective performance measures and the standard deviation of the subjective performance ratings of all employees in a reference group provided by a manager is used as a measure for centrality bias. As the objective performance measures are the same in all treatments of our paper, we only focus on the standard deviation of the subjective performance ratings of all employees in the reference group.

manager.[43] This second variable will be used in supplemental analyses. Higher values for both dependent variables indicate a higher degree of differentiation in the bonus allocation, which means less centrality bias. Furthermore, in the post-experimental questionnaire, managers rated their agreement with the following statement (using a seven-point Likert scale): "*If I would offer a well-performing store manager a much higher bonus than a badly-performing store manager, then my store managers would think that I am fair*". This variable, *Perceived Fairness of Differentiation*, refers to managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations.

Manipulations

In the 2*2 experiment the variables '*information accuracy*' and '*accountability*' were manipulated. We manipulated the variable information accuracy at two levels (high vs. low), as in Bol et al. (2016). Participants in the Low Accuracy conditions received the following message in the case description: "*The performance criteria will give only a moderate picture of the performance of each individual store manager*". Participants in the High Accuracy conditions were instructed as follows: "*The performance criteria will give a fairly accurate picture of the performance of each individual store manager*".

The variable accountability was manipulated at two levels (present vs. absent), in line with Brutus (2010) and Libby et al. (2004). Participants in the Accountability conditions were given the opportunity to provide written comments to justify or clarify their bonus decision to each individual employee separately, whereas the participants in the No Accountability conditions did not get this opportunity. More specifically, participants in the Accountability conditions received additional messages mentioning *"Additionally, you can justify your bonus decision in writing to each individual store manager"* and "*Written justification to the store manager: below you are allowed to indicate to each store manager why you allocate this amount of money to that store manager*".[44] Five text fields for the justification for each individual employee were then provided, next to the five input fields for the bonus amounts.

Participants

We recruited 292 undergraduate students from a business program in a large university. The students received a course credit as an incentive to participate in the experiment. The experiment lasted 30 to 60 minutes (depending on the decision speed of the participants). 129 students were female and 163 were male. They were 20 years old on average (minimum 18 years, maximum 29 years) and they had 8.3 months of work experience, which is significantly different from 0 ($t_{291} = 11.511$; $p < 0.001$).

[43] This measure is similar to the measure for centrality bias in Bol et al. (2016). In this paper centrality bias is captured by calculating the difference in bonus allocated to the strongest performer and the weakest performer.
[44] The justification of performance scores to a virtual person instead of an actual, living person is common in this literature (Bol et al., 2016; Libby et al., 2004; Mero & Motowidlo, 1995; Mero et al., 2003).

3.4 Results

Manipulation checks and Descriptive statistics

In the No Accountability conditions the experimental computer program did not allow for a justification or clarification of the bonus allocation. As such, in these experimental treatments accountability was denied. In the Accountability conditions the experimental computer program provided the possibility for a written justification. All participants in the Accountability conditions chose to make use of this possibility and justified their bonus allocation to each store manager. The manipulation of accountability therefore worked. Furthermore, accountability did indeed lead to increased information processing: participants in the Accountability conditions did take significantly more time to make and justify their bonus decisions (mean = 787 seconds) than the time taken by participants in the No Accountability conditions to make their bonus decision (mean = 247 seconds)($F_{1, 290} = 359.131$; $p < 0.001$).[45] Additionally, various items of the post-experimental questionnaire indicate the effects of accountability on participants' rating behavior.[46] Participants in the Accountability conditions had the perception that *"they took the potential reaction of the store managers to their bonus decision"* significantly more into account than the managers in the No Accountability conditions ($F_{1, 290} = 5.228$; $p = 0.023$). Furthermore, participants in the Accountability conditions admitted significantly more that *"providing each store manager with a similar bonus would lead store managers to think that they insufficiently thought about the bonus decision"* ($F_{1, 290} = 3.339$; $p = 0.069$). Participants in the Accountability conditions stated significantly more than participants in the No Accountability conditions that *"a fair bonus division is important"* ($F_{1,290} = 8.391$; $p = 0.004$) and that *"they divided the bonus fairly"* ($F_{1, 290} = 6.830$; $p = 0.009$). Participants in all conditions thought *"one can avoid a potential reaction by the store managers providing a written justification"* (mean of 5.79 on 7), which is significantly higher than the scale midpoint of 4 ($t_{291} = 26.429$; $p < 0.001$).[47] Participants in the Accountability conditions indicated that *"they thoroughly explained the procedures used during the bonus allocation"* (mean of 4.87 on 7), which is significantly higher than the scale midpoint of 4 ($t_{146} = 7.408$; $p < 0.001$).

In order to test our manipulation of performance information accuracy, we included a manipulation check question in the post-experimental questionnaire. Using a seven-point Likert scale, participants rated their agreement with the statement that *"the three performance*

[45] Every p-value mentioned in this paper is a two-sided p-value.

[46] Using a seven-point Likert scale, we collected this data at the end of the experiment in order to avoid that participants' attention was attracted to the presence or absence of the possibility for justification depending on the condition they were in and in order to avoid leading participants to certain response patterns.

[47] Participants in the No Accountability conditions agreed that *"a written justification of the bonus allocation to the store managers would have been useful"* (mean of 6.06 on 7), which is significantly higher than the scale midpoint of 4 ($t_{144} = 23.965$; $p < 0.001$) and they agreed that *"they would have liked to provide a written justification to the store managers"* (mean of 5.74 on 7), which is significantly higher than the scale midpoint of 4 ($t_{144} = 15.061$; $p < 0.001$). Participants in the Accountability conditions agreed that *"the written justification of the bonus allocation to the store managers was useful"* (mean of 6.00 on 7), which is significantly higher than the scale midpoint of 4 ($t_{146} = 18.913$; $p < 0.001$) and that *"they would dislike allocating the bonus without a justification of the bonus allocation"* (mean of 5.46 on 7), which is significantly higher than the scale midpoint of 4 ($t_{146} = 11.202$; $p < 0.001$).

measures provided a highly accurate image of the performance of each individual store manager". Participants judged the accuracy of the performance measures to be significantly higher in the High Accuracy than in the Low Accuracy conditions ($F_{1, 290} = 27.042$; $p < 0.001$). Thus, our manipulations of accountability and performance information accuracy were successful.

Additionally we asked participants to indicate (on a seven-point scale) their agreement with a couple of statements related to the experimental task. Participants indicated they were motivated (mean of 6.37 on 7), which is significantly higher than the scale midpoint of 4 ($t_{291} = 58.825$; $p < 0.001$). The participants said that the instructions were clear (mean of 6.37 on 7), which is significantly higher than the scale midpoint of 4 ($t_{291} = 48.322$; $p < 0.001$). Participants indicated that they had enough time to make their decisions (mean of 6.8 on 7), which is significantly higher than the scale midpoint of 4 ($t_{291} = 99.241$; $p < 0.001$). Furthermore, participants did not think the task was difficult. The mean response (3.70 on 7) is significantly lower than the scale midpoint of 4 ($t_{291} = 3.192$; $p = 0.002$). All participants allocated the highest bonus to manager C and the lowest bonus to manager E, which is consistent with the performance information overview in Table 3.1. Together, these results suggest that participants felt they had sufficient time and were capable to make a reasonable performance evaluation.

Table 3.2 shows the definitions and descriptive statistics (mean and standard deviation) by experimental condition for the main variables. This table shows the correlation matrix for the main variables as well.

Table 3.2 Descriptive Statistics

Panel A - Mean (Standard Deviation)

	Accuracy[a]			
	High		Low	
	Accountability[b]		Accountability[b]	
	Absent	Present	Absent	Present
Dependent Measure	(n = 72)	(n = 74)	(n = 73)	(n = 73)
Differentiation[c]	5,178.17	5,785.73	4,985.38	4,989.18
	(2,085.94)	(1,976.46)	(1,885.73)	(1,687.75)
Bonus Range[d]	15,368.06	17,361.41	14,863.01	14,910.96
	(5,967.04)	(5,633.34)	(5,517.18)	(4,960.50)
Perceived Fairness	5.86	6.31	5.96	5.92
of Differentiation[e]	(1.42)	(0.83)	(0.93)	(1.19)
Maximum Bonus[f]	18,236.11	19,442.49	18,041.10	18,000.00
	(4,263.32)	(4,048.20)	(3,619.73)	(3,106.89)
Minimum Bonus[g]	2,868.06	2,081.08	3,178.08	3,089.04
	(2,264.22)	(2,131.03)	(2,285.78)	(2,323.41)

Panel B - Pearson Correlation Matrix (n = 292)

	Differentiation	Bonus Range	Perceived Fairness of Differentiation
Differentiation	1.000		
Bonus Range	0.990***	1.000	
Perceived Fairness of Differentiation	0.171***	0.176***	1.000

***, ** and * Indicate p-value, 0.01, 0.05, and 0.1 respectively, two-tailed.

[a] *Accuracy* refers to whether the accuracy of the performance information is high or low. Accuracy is coded as 1 for the High Accuracy conditions and 0 for the Low Accuracy conditions.
[b] *Accountability* refers to whether the manager had an opportunity to provide a written justification for the bonus allocation or not. Accountability is coded as 1 for the Accountability conditions and as 0 for the No Accountability conditions.
[c] *Differentiation* refers to the extent of differentiation in the bonus allocation (i.e. the standard deviation of the bonuses given to each of the 5 store managers).
[d] *Bonus Range* refers to the extent of differentiation in the bonus allocation (i.e. the maximum bonus allocated to a store manager minus the minimum bonus allocated to a store manager.).
[e] *Perceived Fairness of Differentiation* reflects the *managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations*. Using a seven-point Likert scale, managers rated their agreement with the following statement *"If I would offer a well-performing store manager a much higher bonus than a badly-performing store manager, then my store managers would think that I am fair."*.
[f] *Maximum Bonus* refers to the maximum bonus allocated to a store manager.
[g] *Minimum Bonus* refers to the minimum bonus allocated to a store manager.

Hypothesis tests

Our first hypothesis predicts a positive effect of information accuracy on differentiation in bonus allocation. In order to test this hypothesis, we conducted an ANOVA analysis with *Differentiation* as the dependent variable and *Accuracy* and *Accountability* as the independent variables (not tabulated). The ANOVA analysis indicates a significant main effect for *Accuracy* ($F_{1, 288} = 4.875$; $p = 0.028$), in line with the informativeness principle (Holmström, 1979) and hypothesis 1. The ANOVA does not show a significant main effect for *Accountability* ($F_{1, 288} = 1.861$; $p = 0.174$), nor a significant interaction of *Accuracy* and *Accountability* ($F_{1, 288} = 1.816$; $p = 0.179$). We additionally conducted a simple effects analysis. This analysis (not tabulated) indicates that *Accountability* leads to more differentiation in bonus allocations when *Accuracy* is high ($F_{1, 288} = 3.676$, $p = 0. 056$), but not when *Accuracy* is low ($F_{1, 288} = 0.000$, $p = 0.990$). Additionally, the simple effects analysis indicates that *Accuracy* leads to more differentiation in bonus allocations when managers are *Accountable* ($F_{1, 288} = 6.364$, $p = 0.012$), but not when managers are not *Accountable* ($F_{1, 288} = 0.368$, $p = 0.545$).[48]

In order to test hypothesis 2 (Accountability will increase managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations, but only when information accuracy is relatively high), we conducted an ANOVA analysis with *Perceived Fairness of Differentiation* as the dependent variable and *Accuracy* and *Accountability* as the independent variables. Consistent with the reasoning above, ANOVA results (not tabulated) indicate a significant interaction of *Accuracy* and *Accountability* on *Perceived Fairness of Differentiation* ($F_{1, 288} = 3.548$, $p = 0.061$). Simple effects analysis (not tabulated) indicates a significant effect of *Accountability* when *Accuracy* is high ($F_{1, 288} = 5.956$, $p = 0.015$) (i.e. when performance information accuracy is high, accountable managers estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations higher (mean = 6.31 on 7) than non-accountable managers (mean = 5.86 on 7)) and no effect of *Accountability* when *Accuracy* is low ($F_{1, 288} = 0.050$, $p = 0.824$) (i.e. when performance information accuracy is low, accountable managers do not estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations higher (mean = 5.92 on 7) than non-accountable managers (mean = 5.96 on 7)). This result supports hypothesis 2. Simple effects analysis further indicates a significant effect of *Accuracy* when *Accountability* is present ($F_{1, 288} = 4.581$, $p = 0.033$) (i.e. when managers are accountable, they estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations higher when information accuracy is high (mean = 6.31 on 7) than when information accuracy is low (mean = 5.92 on 7)). Furthermore, there is no significant effect of *Accuracy* when *Accountability* is absent ($F_{1, 288} = 0.280$, $p = 0.597$) (i.e. when managers are not accountable, they do not estimate employees' procedural fairness perceptions and acceptance

[48] We also conducted an ANOVA analysis with our alternative dependent variable for centrality bias, *Bonus Range*. The results are similar to the results for *Differentiation*. ANOVA analysis on *Bonus Range* (not tabulated) shows a significant main effect for *Accuracy* ($F_{1, 288} = 5.212$; $p = 0.023$), no significant main effect for *Accountability* ($F_{1, 288} = 2.486$; $p = 0.116$), nor a significant interaction effect of *Accuracy* and *Accountability* on *Bonus Range* ($F_{1, 288} = 2.258$; $p = 0.134$). Simple effects analysis indicates that *Accountability* leads to differentiation in bonus allocations when *Accuracy* is high ($F_{1, 288} = 4.741$, $p = 0.030$), but not when *Accuracy* is low ($F_{1, 288} = 0.003$, $p = 0.958$). *Accuracy* leads to more differentiation in bonus allocations when managers are *Accountable* ($F_{1, 288} = 7.215$, $p = 0.008$), but not when managers are not *Accountable* ($F_{1, 288} = 0.302$, $p = 0.583$).

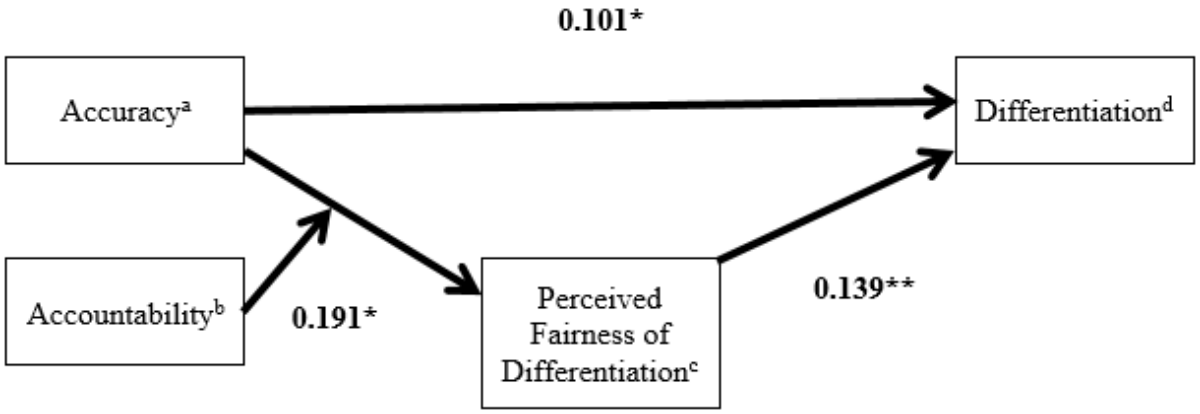
of differentiation in bonus allocations higher when information accuracy is high (mean = 5.86 on 7) than when information accuracy is low (mean = 5.96 on 7)).

To further investigate our expectations, we conducted a path analysis to test the overall process and the specific paths by which *Accuracy* and *Accountability* affect *Perceived Fairness of Differentiation* and eventually *Differentiation* in the bonus allocations (see also Barton & Mercer, 2005; Christ et al., 2012; Masschelein et al., 2012). Figure 3.2 displays the path model with standardized path coefficients significant at the 10% level or less (two-tailed). The model is suitable for the data, as indicated by several tests of goodness of fit. The Bentler Comparative Fit Index (CFI) (0.9560) indicates that our model is significantly better than the null model, since its value is above the recommended minimum value of 0.95 (Byrne 2001). The Goodness of Fit Index (GFI) (0.9798) indicates a good fit, since its value is above the generally accepted minimum value of 0.95 (Byrne 2001). However, χ^2 (df = 2, n = 292) = 16.7425 (p = 0.0002) indicates that the model might be less suitable for the data, though this test is very sensitive to large sample sizes. The results of the path model indicate that *Accuracy* leads directly to more differentiation in bonus allocations (0.101; $p < 0.1$), confirming hypothesis 1. Furthermore, a significant interaction between *Accuracy* and *Accountability* indicates that managers estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations higher when they are accountable and have accurate information at their disposal (0.191; $p < 0.1$), confirming hypothesis 2. Finally, managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations are positively related to the extent to which managers differentiate in their bonus allocations (0.139; $p < 0.05$), confirming hypothesis 3. Overall, our findings confirm the theoretical reasoning above.

We conducted a similar path analysis to test the overall process for the alternative dependent variable *Bonus Range* as well.[49] Results (not tabulated) are consistent. *Accuracy* leads directly to more differentiation in bonus allocations (0.100; $p < 0.1$), confirming hypothesis 1. A significant interaction between *Accuracy* and *Accountability* indicates that managers estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations higher when they are accountable and have accurate information at their disposal (0.191; $p < 0.1$), confirming hypothesis 2. Managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations are positively related to the extent to which managers differentiate in their bonus allocations (0.137; $p < 0.05$), confirming hypothesis 3.

[49] The Goodness of Fit Index (GFI) (0.9699) indicates a good fit, since its value is above the generally accepted minimum value of 0.95 (Byrne 2001). However, χ^2 (df = 2, n = 292) = 26.0549 ($p < 0.0001$) indicates that the model might be less suitable for the data, though this test is very sensitive to large sample sizes.

Figure 3.2 Path Analysis Results



***, ** and * Indicate p-value, 0.01, 0.05, and 0.1 respectively, two-tailed.

[a] *Accuracy* refers to whether the accuracy of the performance information is high or low. Accuracy is coded as 1 for the High Accuracy conditions and 0 for the Low Accuracy conditions.

[b] *Accountability* refers to whether the manager had an opportunity to provide a written justification for the bonus allocation or not. Accountability is coded as 1 for the Accountability conditions and as 0 for the No Accountability conditions.

[c] *Perceived Fairness of Differentiation* reflects the *managers' estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations*. Using a seven-point Likert scale, managers rated their agreement with the following statement *"If I would offer a well-performing store manager a much higher bonus than a badly-performing store manager, then my store managers would think that I am fair".*

[d] *Differentiation* refers to the extent of differentiation in the bonus allocation (i.e. the standard deviation of the bonuses given to each of the 5 store managers).

Fit Indices: χ^2 (df = 2, n = 292) = 16.7425 (p = 0.0002); Goodness of Fit Index (GFI) = 0.9798; Adjusted GFI (AGFI) = 0.8487; Bentler Comparative Fit Index = 0.9560; standardized RMR = 0.1001
Standardized path coefficients are presented.

Additional analyses

Prior literature on compression in evaluations indicates that the compression in evaluations is caused by an asymmetric effect i.e. managers increase the bonus of weaker performers in order to minimize personal costs such as confrontation costs, while maintaining the same levels of bonuses for stronger performers (Bol, 2011; Bol et al., 2016). These studies examine the context of independent bonuses, in which they find that bonuses allocated to weaker performers do not affect bonuses allocated to stronger performers. Since in our setting, however, the bonus pool is fixed, the dynamics of managers' compressed rating decisions might be different (Bol et al., 2016). A fixed bonus pool implies an inherent interdependence between the bonuses allocated to all employees. In other words, it is impossible to favor a weaker performer without affecting one or more other employees. We expect managers to allocate bonuses symmetrically, in the sense that they might offer weaker performers higher bonuses than they deserve and stronger performers lower bonuses than they deserve. In the following paragraph, we investigate this expectation empirically.

If we examine the variables *Maximum Bonus,* the highest bonus allocated by the manager, and *Minimum Bonus*, the lowest bonus allocated by the manager, in closer detail, we see that they almost mirror each other. ANOVA analysis with *Maximum Bonus* as the dependent

variable and *Accuracy* and *Accountability* as the independent variables (not tabulated) indicates a significant main effect for *Accuracy* ($F_{1, 288} = 3.416$; $p = 0.066$): more accurate information leads to a higher *Maximum Bonus* in line with the informativeness principle (Holmström, 1979) and hypothesis 1. The ANOVA does not show a significant main effect for *Accountability* ($F_{1, 288} = 1.730$; $p = 0.189$), nor a significant interaction of *Accuracy* and *Accountability* ($F_{1, 288} = 1.983$; $p = 0.160$). We additionally conducted a simple effects analysis. This analysis (not tabulated) indicates that *Accountability* leads to a higher *Maximum Bonus* when *Accuracy* is high ($F_{1, 288} = 3.708$, $p = 0.055$), but not when *Accuracy* is low ($F_{1, 288} = 0.004$, $p = 0.948$). Additionally, the simple effects analysis indicates that *Accuracy* leads to a higher *Maximum Bonus* when managers are *Accountable* ($F_{1, 288} = 5.338$, $p = 0.022$), but not when managers are not *Accountable* ($F_{1, 288} = 0.096$, $p = 0.757$). ANOVA analysis with *Minimum Bonus* as the dependent variable and *Accuracy* and *Accountability* as the independent variables (not tabulated) indicates a significant main effect for *Accuracy* ($F_{1, 288} = 6.251$; $p = 0.013$). In other words, more accurate information leads to a lower *Minimum Bonus* in line with the informativeness principle (Holmström, 1979) and hypothesis 1. The ANOVA does not show a significant main effect for *Accountability* ($F_{1, 288} = 2.762$; $p = 0.098$), nor a significant interaction of *Accuracy* and *Accountability* ($F_{1, 288} = 1.753$; $p = 0.187$). We additionally conducted a simple effects analysis. This analysis (not tabulated) indicates that *Accountability* leads to a lower *Minimum Bonus* when *Accuracy* is high ($F_{1, 288} = 4.457$, $p = 0.036$), but not when *Accuracy* is low ($F_{1, 288} = 0.057$, $p = 0.811$). Additionally, the simple effects analysis indicates that *Accuracy* leads to a lower *Minimum Bonus* when managers are *Accountable* ($F_{1, 288} = 7.363$, $p = 0.007$), but not when managers are not *Accountable* ($F_{1, 288} = 0.687$, $p = 0.408$). It means that the observed pattern of managers' differentiation in the bonus allocation is symmetrically caused both by a higher bonus allocated to the strongest store manager and by a lower bonus allocated to the weakest store manager.

We further look at the bonuses allocated to the three mediocre store managers. ANOVA analysis with the average bonus allocated to the three mediocre store managers as the dependent variable and *Accuracy* and *Accountability* as the independent variables (not tabulated) indicates an insignificant main effect for *Accuracy* ($F_{1, 288} = 0.702$; $p = 0.403$), an insignificant main effect for *Accountability* ($F_{1, 288} = 0.607$; $p = 0.437$) and an insignificant interaction of *Accuracy* and *Accountability* ($F_{1, 288} = 0.345$; $p = 0.557$). ANOVA analysis with the standard deviation of the bonuses allocated to the three mediocre store managers as the dependent variable and *Accuracy* and *Accountability* as the independent variables (not tabulated) indicates an insignificant main effect for *Accuracy* ($F_{1, 288} = 0.047$; $p = 0.828$), an insignificant main effect for *Accountability* ($F_{1, 288} = 0.455$; $p = 0.500$) and an insignificant interaction of *Accuracy* and *Accountability* ($F_{1, 288} = 0.027$; $p = 0.870$). As such, the average level of the bonus allocated to mediocre store managers and the compression in the bonuses of mediocre store managers is not affected by information accuracy nor process accountability. It means that the observed pattern of managers' differentiation in the bonus allocation is caused both by a higher bonus allocated to the strongest store manager and by a lower bonus allocated to the weakest store manager, but not by adjustments to the bonuses of mediocre store managers.

3.5 Conclusion and discussion

In this paper, we investigate the common practice that allows managers subjectively evaluate and reward employees based on a number of key performance measures (Gibbs et al., 2004; Golman & Bhatia, 2012; Rajan & Reichelstein, 2006). One of the main problems with subjective performance evaluation is that managers usually do not differentiate adequately between employees (Bol, 2008; Moers, 2005). This is not desirable for organizations, as more differentiation leads to favorable organizational outcomes such as higher employee incentives, productivity, performance and performance improvement and better recognition of which employees might be eligible for promotion or additional training (Ahn et al., 2010; Baker et al., 1988; Bol, 2011; Golman & Bhatia, 2012).

We develop and find support for a causal model that explains how managers subjectively allocate a fixed bonus pool to a number of employees. Specifically, our model shows how (the perceived) accuracy of the performance information and a manager's opportunity to write out a justification for the bonus allocation to her employees (process accountability) influence how much a manager differentiates her bonus allocations amongst employees. Higher (perceived) information accuracy directly leads to more differentiation in bonus allocations compared to lower (perceived) information accuracy, supporting the informativeness principle (Holmström, 1979). Furthermore, (perceived) information accuracy and the possibility to write out a justification for the bonus allocation jointly influence the manager's estimates of employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations. More specifically, managers estimate employees' procedural fairness perceptions and acceptance of differentiation in bonus allocations as higher when the performance information accuracy is high and when they are allowed to justify the bonus allocation to the employees. Managers will consequently differentiate more in their bonus allocations to employees, in line with their estimates of employees' perceived procedural fairness.

Allowing managers to write out a justification for their bonus allocation provokes accountability urges that lead them to process the information better and to take better decisions (Ashton, 1990; Gibbins & Newton, 1994; Kennedy, 1993; Lerner & Tetlock, 1999; Mero & Motowidlo, 1995; Mero et al., 2003), but the manager will only use data that is perceived fit for use, of sufficiently high quality (Pipino et al., 2002; Wang & Strong, 1996). Organizations should therefore invest in either higher data quality, or either perceptions of higher data quality. Since perceptions of data quality (information accuracy) vary across different users of the performance information (Pipino et al., 2002; Wang & Strong, 1996), organizations could consider dedicating time and resources to specific communication aimed to alter or improve managers' perceptions about the data quality of the performance measures (Libby et al., 2004). Such perceptions will eventually drive the use of the performance measures in performance evaluation. More accurate information often comes at a cost and firms need to make a trade-off between higher data quality and costs for investments in data-gathering and higher data quality (Bol, 2011). Based on our observations, we conclude that extra investments in higher (perceptions of) data quality (and the often extensive related costs) are worthwhile, especially when organizations provide managers the possibility to justify their bonus allocations: it will encourage managers to spend time and effort on information processing and will strengthen the pay-for-performance relationship. Our results are in line with the call for dedicating enough

time, effort and/or resources to the performance information gathering process by Bol (2011), because we observe that extra time and effort spent on information processing during the performance evaluation (induced by accountability) cannot make up for a lack of sufficiently accurate performance information. Allowing a manager to justify her performance evaluation decision when she perceives the performance information as inaccurate is not worthwhile, as it is a waste of the manager's valuable productive time without altering her estimates of employees' procedural fairness perceptions, or, at the bottom line, the degree to which she differentiates in the bonus allocation. In our experiment, for instance, accountable managers with inaccurate performance information spent on average 828 seconds on the performance evaluation and justification, which is significantly higher than the 748 seconds spent by the accountable managers with accurate performance information.

This study contributes to the accounting and management literature on subjective performance evaluation based on scorecards (Banker et al., 2004; Cardinaels & van Veen-Dirks, 2010; Humphreys & Trotman, 2011; Ittner et al., 2003; Lipe & Salterio, 2000; Libby et al., 2004). Our findings are useful, as they provide interesting insights on managers' rating behavior, and more specifically centrality bias in performance evaluations (Ahn et al., 2010; Bol, 2011; Bol et al., 2016; Golman & Bhatia, 2011; Levin, 2003; MacLeod, 2003; Moers, 2005; Prendergast & Topel, 1993). First, we contribute to the literature indicating that choices related to the design of the management control system should not be made in isolation, as they influence evaluation bias (Bol et al., 2016; Libby et al., 2004). We research a similar setting as in Bol et al. (2016) and our results are strongly in line with theirs. However, while both papers are related to information asymmetry from the perspective of the employees, process accountability and outcome transparency are different concepts (Castilla, 2008, 2015).[50] Outcome transparency (the focus of Bol et al., 2016) relates to a forced disclosure of the outcomes of the bonus allocation to all employees, whereas process accountability (the focus of this paper) relates to a voluntary disclosure of the bonus allocation process to each individual employee separately, but not to the disclosure of outcomes of the bonus allocation. Pay secrecy is an established organizational practice that firms often do not want to alter, because it reduces or avoids conflicts in the workplace (Colella et al., 2007). In this context, process accountability might be a valid alternative to outcome transparency that is more easy to install and leads to similar results for managers' performance evaluation.

Second, we provide more details about the underlying process leading to compressed rating behavior, by examining how the two control system design elements, information accuracy and process accountability, jointly affect managers' estimates of employees' procedural fairness perceptions and consequently managers' performance evaluation (Bol, 2011; Bol et al., 2016; Moers, 2005). Third, we contribute to this literature by investigating the allocation of a fixed bonus pool. Prior literature focuses on the allocation of independent bonuses, showing that managers tend to rate asymmetrically. Compression in bonus allocations is an indirect effect of managers' decision to inflate weaker employees' bonuses (Bol, 2011; Bol et al., 2016; Moers, 2005). However, with a fixed bonus pool managers can no longer inflate weaker

[50] See Castilla (2008, 2015) for a discussion of the differences and similarities between transparency and process acountability.

employees' bonuses without affecting other employees' bonuses and risking confrontation with these employees. Therefore the allocation of a fixed bonus pool involves a deliberate decision of a manager to differentiate in bonus allocation, taking into account the fairness for all employees (Walster et al., 1973). Balancing between employees in the bonus allocation is inherent given the zero-sum nature of fixed bonus pools (Baiman & Rajan, 1995; Rajan & Reichelstein, 2006) and this inherent interdependence might change the dynamics of managers (Bol et al., 2016). We demonstrate indeed that the dynamics of managers' rating decisions change when using a fixed bonus pool instead of independent bonuses. Firstly, our research shows that manager's decision to compress bonus allocations with a fixed bonus pool is symmetric i.e. managers tend to inflate bonuses of weaker employees as well as to decrease bonuses of stronger employees. Secondly, our research indicates that the manager's decision to compress bonus allocations with a fixed bonus pool is directly dependent on performance information accuracy. Third, we add to the literature on accountability and its debiasing potential in performance evaluations as well (Brutus, 2010; Hoffman & Patton, 1997; Kennedy, 1993; Libby et al., 2004; Mero et al., 2003; Mero et al., 2007; Roch, 2005; Shore & Tashchian, 2002). Some judgment biases are more effort-related than others and hence some judgment biases are more susceptible to accountability's debiasing potential (Kennedy, 1993). We demonstrate that accountability's debiasing potential for the centrality bias depends on the accuracy of the performance information.

We recognize that this study is subject to a number of limitations. Firstly, participants in the computer-based experimental task evaluated a hypothetical case and did not actively interact with the people they had to evaluate. This setting did not allow for potential conflict with one of the subordinates and limited the fear for negative reactions by one of the subordinates. The absence of physical interaction is highly common in the literature (Bol et al., 2016; Klimoski & Inks, 1990; Mero et al., 2007; Shapiro, 1975). As our results are already strong in a situation without a potential conflict, we are convinced that actual, face-to-face interaction would make the results even more significant. We assume that in a situation with actual face-to-face interaction and more being at stake, managers in the Low Accuracy - Accountability condition might have attempted to avoid confrontation with their subordinates and therefore might have allocated less differentiated bonuses than participants in the Low Accuracy – No Accountability condition. The differentiation is expected to be even lower when accountable managers would find out, through their extra information-processing, that the performance information is so inaccurate that a differentiated bonus allocation is impossible to justify to the employees, whereas managers that are not accountable would not come to this conclusion. However, we consider it unlikely that organizations would use performance information in the performance evaluation that is so inaccurate that accountable managers consider it worthless to use. Additionally, we consider it unlikely that managers that are not accountable would not be able to detect the inferiority of the 'worthless' performance information. We leave this for future research. Furthermore, we operationalized accountability as a 'possibility' to write out a justification to the subordinate, while it could be a requirement (Libby et al., 2004; Mero & Motowidlo, 1995; Mero et al., 2007). We believe that our results would have been even more outspoken if we would have used a requirement instead of a possibility. Realistic alternatives are to introduce a justification towards a superior, or towards both a superior and the

subordinates (Ferris et al., 2008; Libby et al., 2004; Mero et al., 2007). The justification of performance evaluations to subordinates is the norm (Libby et al., 2004), but we leave it for future research to investigate whether the justification towards a superior or towards both a superior and the subordinates affects our results. Lastly, the performance measures on which the participants had to base their performance evaluations in our experiment were all expressed in terms of a numerical, quantitative rating scale. This may have affected participants' perceptions of the information accuracy. More specifically, participants may perceive numerical performance measures as more accurate and objective than measures rated on a verbal scale or more subjective performance measures such as perceptions of the supervisor regarding the behavior of her subordinates.[51] Quantitative performance measures are often called objective. However, the measures can often be manipulated by the employee or the manager and they may often not measure exactly what they are thought to measure, so this terminology is misleading (Gibbs et al., 2004). If participants in the Low Accuracy conditions thought the performance measures were rather accurate and objective, this may have reduced the strength of the manipulation of *Information Accuracy.* Nonetheless, we obtained significant results, but we recommend future research to include more subjective, non-numerical performance measures.

[51] Even in the Low Accuracy conditions participants thought the performance measures were quite accurate. Their average response to the statement *"the three performance measures provided a highly accurate image of the performance of each individual store manager"* was 4.13 on 7, which is not significantly different from the scale midpoint of 4 ($t_{145} = 1.089$; $p = 0.278$), but still rather high.

Chapter 4
Performance evaluation in a multitasking environment: the effect of task complexity, distorted time allocation and detailed relative performance information

Abstract
Employees often work on multiple tasks, involving the need to schedule these tasks appropriately (Gonzalez & Mark, 2004). Previous research indicates that employees often distort their effort allocation across tasks away from the firm-preferred effort allocation as a reaction to relative performance information (RPI) (Hannan et al., 2013). This distortion reduces employee performance (Hannan et al., 2013) and ultimately firm performance (Farkas, 2013). In this paper we describe the results of two experiments in which two employees work on two tasks for a manager. In both experiments, we manipulated task complexity (a context with two complex tasks versus a context with two less complex tasks). The paper examines the joint role of task complexity and distortion in effort allocation across tasks on employee performance. Additionally, we investigate how a management control system providing detailed RPI and a discretionary bonus affects employees' effort allocation across tasks. Our findings indicate that a distorted allocation of effort across tasks indeed leads to lower employee overall performance on both tasks, consistent with Hannan et al. (2013). However, the effect of distorted allocation of effort across tasks on overall employee performance on both tasks is less negative in a work environment consisting of complex tasks than in a work environment with simple tasks. Furthermore, we find that, in the first period, employees will focus on the task for which they have the highest skills or find easiest in order to perform at least well on that task. However, managers' bonus allocation is in line with the RPI they receive about the employees. Consistent with social comparison theory and expectancy theory (Festinger, 1954; Vroom, 1964), we find that after the provision of RPI and the performance evaluation, employees reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague.

Keywords: task complexity, detailed relative performance information, multitasking, social comparison, performance evaluation

4.1 Introduction

This paper investigates how employees' effort allocation across tasks and task complexity[52] jointly affect employees' overall performance in a multitasking environment. Additionally, we research how employees' effort allocations across tasks are influenced by managers' performance evaluations based on detailed relative performance information (RPI). [53] Nowadays, time management has become an important aspect of employees' job. Employees typically engage in multiple tasks, involving the need to schedule these tasks appropriately (Adler & Benbunan-Fich, 2012; Farkas, 2013; Gonzalez & Mark, 2004; Kushleyeva et al., 2005; Payne et al., 2007). A trade-off in attention allocation exists as effort expended on one task cannot be expended on another task (Eichhorn, 2016; Farkas, 2013). Employees often overly focus on some tasks, neglecting the other ones, leading to overachievement on some tasks and underachievement on others (Baker et al., 1994; Hannan et al., 2013; Ittner et al., 2003). Such dysfunctional employee behavior is unwanted, as employees that distort their effort allocations across tasks away from the firm-preferred allocations reduce employee overall performance, firm productivity and profits (Farkas, 2013; Hannan et al., 2013).

Organizations can therefore assist their employees by designing management control systems that motivate the appropriate level and allocation of effort across tasks according to firm preferences (Chenhall, 2003; Farkas, 2013; Hannan et al., 2013). One possibility is to work with financial incentives based on detailed relative performance information to redirect employees' effort allocations across tasks (Brüggen, 2011; Brüggen & Moers, 2007; Eichhorn, 2016; Farkas, 2013; Hannan et al., 2013). Prior research demonstrated the positive effect of RPI on employee performance in a single-task setting (Kramer et al., 2016; Murthy & Schafer, 2011; Tafkov, 2013) and concluded that this positive effect is larger under a performance-based contract than under a flat-wage contract (Tafkov, 2013). However, when RPI is provided in a multi-tasking setting without financial incentives, it reduces performance and does not lead to firm-preferred effort allocations (Hannan et al., 2013). We are therefore interested in investigating whether tying financial incentives to RPI would be useful to reach firm-preferred effort allocations and better performance in a multi-task setting. An earlier study on financial incentives in a multi-tasking setting (Farkas, 2013) demonstrates that financial incentives are not effective at directing employee effort or improving performance. She also demonstrates that RPI has a negative effect on employees' effort allocation and performance in the presence of financial incentives. However, this result may be due to the fact that employees did not believe they were able to obtain the compensation (Farkas, 2013). In contrast, we examine how employees would act when they are evaluated under a discretionary bonus pool that brings the bonus into reach. In a multitasking environment, a discretionary system in which the manager has the final say in the evaluation seems particularly useful, since it is impossible to design and install a complete contract specifying all possible outcomes and related employee rewards in this context (Gibbs et al., 2004; MacLeod and Parent, 1999; Rajan & Reichelstein, 2006).

[52] *"Broadly defined, task complexity refers to the amount of attention or processing a task requires as well as the amount of structure and clarity the task provides"* (Bonner & Sprinkle, 2002; Campbell, 1988; Wood, 1986).

[53] Relative performance information is information regarding some aspect of peer performance (Farkas, 2013). Detailed relative performance information is information about the specific performance levels of peers (Kramer et al., 2016).

Additionally, discretionary bonus pools mitigate the occurrence of distorted performance, which is often associated with formula-based performance evaluations (Ittner et al., 2003).

In this paper, we first investigate how employees' distortion in effort allocation across tasks away from the firm-preferred equal effort allocation and task complexity jointly affect employees' overall performance (hypothesis 1 and 2). We research employees' effort allocation and its effect on their overall performance across settings with different levels of task complexity (a setting with complex tasks versus a setting with less complex tasks), because accounting task settings can vary strongly in complexity and complexity is one of the most important determinants of performance (Bonner & Sprinkle, 2002). Our first hypothesis predicts that employees' overall performance is lower when effort allocations are distorted (Hannan et al., 2013). Furthermore, the extant literature also highlights the role of task complexity reducing employee performance because employees' abilities increasingly cannot meet task complexity (Bonner et al., 2000; Bonner & Sprinkle, 2002; Bonner, 2008, p. 161). However, since complex tasks require both additional skills and additional effort from employees (Bonner & Sprinkle, 2002), we predict in hypothesis 2 that in a setting with complex tasks a distorted effort allocation that overly focuses on a single task and therefore allocates additional effort to that task may be beneficial for solving that single task and that thus the negative effect of distorted allocation of effort across tasks on employee overall performance is less negative in a context of high task complexity than in a context of low task complexity.

Second, we research how providing detailed RPI affects employees' distortion of effort allocation across tasks away from the firm-preferred allocations (Hypothesis 3a, 3b and 5) and managers' allocation of the discretionary bonus (hypothesis 4a and 4b). Employees typically spend more time on tasks for which they have higher skills (hypothesis 3a) or on easier, more productive tasks (hypothesis 3b) in order to maximize their return to effort (Duggan et al., 2013; Farkas, 2013; Janssen et al., 2011; Payne et al., 2007). By doing so, employees may distort their effort allocations across tasks away from firm-preferred effort allocations. However, a management control system providing detailed RPI and installing a performance evaluation and reward system containing bonuses, can steer employees towards firms objectives and preferences (Chenhall, 2003). Based on equity theory, we predict that managers will allocate their bonus in line with the RPI (hypothesis 4a) (Adams, 1963, 1965; Walster et al., 1973) and that managers will punish deviations from firm-preferred effort allocations, in line with the objectives of the firm (hypothesis 4b) (Chenhall, 2003). Next, the provision of detailed relative performance information in the performance evaluation process allows employees to compare themselves with their colleagues and this social comparison will affect employees' effort allocation across tasks (Festinger, 1954; Hannan et al., 2013). However, social comparison theory and self-affirmation theory provide contradicting predictions related to social comparison (Eichhorn, 2016; Festinger, 1954; Steele, 1988). According to social comparison theory, social comparison leads to competition and a drive to outperform colleagues on all tasks (Eichhorn, 2016; Festinger, 1954, P. 124 & 126; Tafkov, 2013). Therefore, in line with social comparison theory, we predict that employees working for multiple time periods, will in the next period focus more on the tasks they neglected before, and less on the task they overly focused on earlier (hypothesis 5). In contrast, self-affirmation theory predicts that social comparison causes employees to overly focus on some tasks at the

expense of other tasks, because they want to counter a threat to their self-image in one area by affirming their competence in another area (Steele, 1988) and/or satisfy an innate desire for social distinction (Frey, 2007) by outperforming their peers in at least one area (Farkas, 2013; Hannan et al., 2013). Therefore, self-affirmation theory predicts that employees will focus more on the tasks they overly focused on earlier, and focus less on the task they neglected before (Hannan et al., 2013). However, managers can use the management control system to induce an appropriate allocation of effort across tasks according to firm preferences (hypothesis 4a, 4b & 5) (Chenhall, 2003; Farkas, 2013; Hannan et al., 2013). Employees will assess the likelihood that increased effort will yield higher performance and thus higher rewards and will allocate their effort accordingly (Farkas, 2013; Vroom, 1964). The discretionary bonus attached to performance on both tasks limits the opportunity for self-affirmation in a single area of performance and reduces the occurrence of disruptive employee behavior. Installing a performance evaluation system based on detailed RPI allows managers to thoroughly analyze employees' performance, to detect and punish employees' distorted performance, and to mitigate dysfunctional behavior in the future (Baker et al., 1994; Ittner et al., 2003). Therefore the performance evaluation system can incentivize employees to allocate their effort across tasks in line with firm preferences. The detailed RPI indicates precisely how employees should shift their time allocation in order to perform the tasks better than their colleague. Hypothesis 5 predicts that employees will shift their time allocation from the task on which they outperformed their colleague - because they overly focused on that task or because they had higher abilities for that task - to the task for which they underperformed their colleague - because they neglected that task or had lower abilities for that task (Festinger, 1954; Vroom, 1964). We expect that by doing so employees allocate their effort more in line with firm preferences.

We address our research questions by means of two experiments in which two employees work for a manager in a multitasking (dual-task) environment for two periods. The employees can freely decide when to switch between tasks and how much time to allocate to each of two independent tasks. In both experiments, we manipulate the complexity of the tasks that the employees performed (*task complexity*; low vs. high) such that employees work on two simple tasks or two complex tasks. Afterwards, the managers receive a set of performance measures per employee on each task, they reward the employees with a bonus and they can motivate their evaluation towards the employees. The two experiments differ in terms of the management control system. In the first experiment the managers are required to allocate one bonus pool for overall performance on the two tasks. In the second experiment managers can allocate a separate bonus pool for performance on each task separately.

Consistent with our predictions we find that a distorted allocation of effort across tasks leads to lower employee overall performance on both tasks (hypothesis 1) and that the negative effect of distorted allocation of effort across tasks on overall employee performance on both tasks is less negative for complex tasks than for simple tasks (hypothesis 2). Further, our findings indicate that, in the first period, employees will focus on the task for which they have the highest skills (hypothesis 3a), employees will focus on the easiest task in order to perform at least well on that task (hypothesis 3b) and after the provision of RPI and the performance evaluation, employees reallocate their time across task such that they focus more (less) on tasks

for which they under(out)performed relative to their colleague (hypothesis 5). Finally, we confirm hypothesis 4a that managers' bonus allocation is in line with the RPI, but we do not find evidence for hypothesis 4b that, ceteris paribus, managers' bonus allocation will be lower when employees' performance deviates from the firm-preferred effort allocation.

With this research we contribute to the growing stream of accounting research investigating effort allocation concerns in multi-task environments (Bonner & Sprinkle, 2002; Brüggen & Moers, 2007; Hannan et al., 2013, 2017; Hecht et al., 2012). Especially, research on performance evaluations in a multitasking setting is lacking (Bailey & Konstan, 2006; Buser & Peter, 2012; Dux et al., 2009; Monsell, 2003; Ren et al., 2009; Rubinstein et al., 2001; Stoet et al., 2013). First, we provide additional insights on how performance-contingent monetary incentives and detailed RPI can affect employees' allocations and levels of effort among an employee's various responsibilities in small groups (Bonner & Sprinkle, 2002; Hannan et al., 2013, 2017). Most existing RPI studies provided employees with information about their relative rank instead of detailed RPI information.[54] However, only detailed information can indicate employees exactly how their performance on the different tasks compares to the performance of their peers and how they could change their effort level and allocation (Kramer et al., 2016). Furthermore, many studies do not provide any performance-dependent pay [55], but performance-based incentives are omnipresent and affect employee performance (Bonner & Sprinkle, 2002; Eichhorn, 2016). Lastly, the majority of studies focuses on groups of 5 peers (Hannan et al., 2013; Kramer et al., 2016), yet, social comparison concerns and competitive behavior amongst peers as well as employee motivation and effort decrease if the number of peers increases (Garcia & Tor, 2009; Eichhorn, 2016).

Second, we show the extent to which this effort allocation across tasks affects individual effort and task performance depending on task complexity (Bonner & Sprinkle, 2002; Hannan et al., 2013). Our study thereby aims to respond to the call for more research on how the negative effects of task complexity on performance can be attenuated (Bonner, 2008, p. 167; Gillie & Broadbent, 1989; Ren et al., 2009; Rubinstein et al., 2001; Segal, 2004). A potential important implication for practice is that, in a working environment with low task complexity, providing detailed RPI and basing bonus allocations to this information are efficient means to increase employee performance. In contrast, in a working environment with high task complexity, detailed RPI and bonus allocations attached to detailed RPI improve employee performance to a smaller extent. Lastly, RPI and bonus allocations attached to detailed RPI are able to steer employees effort allocations across tasks towards firm preferences.

The remainder of this paper is structured as follows. The next section presents the relevant literature and the development of the hypotheses. Section III discusses the experimental design and section IV focuses on the results. In section V we provide the discussion and conclusion.

[54] Some studies on RPI in multitasking environments only provide information about the relative rank (Farkas, 2013; Hannan et al., 2013) others provide detailed RPI (Hannan et al., 2017) and other studies on RPI in single-task environments provide both relative ranks and detailed RPI (Kramer et al., 2016).

[55] Some studies introduce financial incentives (Brüggen, 2011; Christ et al., 2016; Eichhorn, 2016; Farkas, 2013; Guymon, 2008) others do not (Hannan et al., 2013; Kramer et al., 2016).

4.2 Literature & Hypotheses

In a first step, we research how the impact of employees' effort (re)allocations across tasks on overall employee performance depends on the complexity of the tasks at hand. In a second step, we research how the provision of detailed RPI and bonus allocations based on RPI affect employees' effort allocations across tasks.

Distorted effort allocation across tasks, task complexity and employee overall performance

The effort allocation decision across tasks is an important decision when working in a multi-tasking context.[56, 57] Employees have to decide how much time to allocate to each task and time allocated to one task cannot be allocated to another task (Eichhorn, 2016; Farkas, 2013; Hannan et al., 2013). The effort allocation can have important consequences for employee performance and firm performance when the returns for the various tasks differ (Farkas, 2013; Hannan et al., 2013), but often the completion of less rewarding tasks is important or required as well (Farkas, 2013). Firms often want their employees to focus on several tasks at the same time; e.g. information workers perform administrative deskwork, respond to e-mails and phone calls or participate in scheduled and unscheduled meetings (Gonzalez & Mark, 2004) or employees in a customer service department focus on cross selling or up selling while providing excellent service to customers (Jasmand et al., 2012) or professors supervise students, perform departmental duties and participate in different research projects (Gonzalez & Mark, 2004). Firms often desire a certain allocation of effort across tasks (Farkas, 2013). The distortion of effort across tasks away from the firm-preferred allocation[58] decreases employees' overall performance when different tasks have (similar) diminishing marginal returns to effort, a feature inherent to many tasks (Hannan et al., 2013). Additionally, this distortion leads to a decrease in the firm's overall productivity and profit (Farkas, 2013).

[56] Multitasking is the result of both interruptions caused by an external party as well as internal choices to switch tasks (Adler & Benbunan-Fich, 2012). Externally caused interruptions and their effects on task performance for tasks of different complexity have been extensively researched (Bailey & Konstan, 2006; Borst et al., 2010; Dux et al., 2009; Gillie & Broadbent, 1989; Ren et al., 2009; Rubinstein et al., 2001; Segal, 2004; Speier et al., 1999; Speier et al., 2003; Zijlstra et al., 1999), but empirical literature on self-imposed interruptions is sparse (Adler & Benbunan-Fich, 2012). However, internal interruptions occur as often as external interruptions (Gonzalez & Mark, 2004), so further research on the effects of internal interruptions on performance is warranted (Adler & Benbunan-Fich, 2012). Therefore, we unravel employees' multitasking behavior under discretionary multitasking i.e. self-imposed, internal choices to switch tasks.

[57] Note that our definition of multitasking differs from the research of Adler & Benbunan-Fich (2012), Buser & Peter (2012) and Madjar & Shalley (2008). They do not allow for the reallocation of time between tasks; they keep the total time that can be allocated to each separate task constant. They thereby ignore a possible advantage of working on multiple tasks concurrently, the possibility of time re-allocation and better time-management (König, et al., 2005). This is in contrast to everyday life in which people typically have to deal with different, independent tasks and need to schedule these tasks appropriately in order to obtain their goals within the limited time available (time management) (Payne et al., 2007). Many real-world complex tasks involve multitasking in which performers themselves decide when to switch between tasks and how much time to devote to each task (Kushleyeva et al., 2005).

[58] Firm-preferred effort allocations are employee effort allocations in line with the organization's desires (Farkas, 2013). In our experiments, the firm prefers an equal allocation of effort across tasks, in line with Hannan et al. (2013).

However, the relationship between distorted allocation of effort across tasks and overall employee performance depends on task complexity as well. Complexity refers to the amount of attention or processing a task requires as well as the amount of structure and clarity the task provides (Bonner & Sprinkle, 2002; Campbell, 1988; Wood, 1986). Complex tasks require additional skills, additional effort (Bonner & Sprinkle, 2002) and a better task strategy (Farkas, 2013). When performing complex tasks, employees in general lose intrinsic and extrinsic motivation[59], because they do not feel competent for the job (i.e. they lack ability for the job) and therefore believe the effort will not yield a reward (Bonner et al., 2000; Bonner & Sprinkle, 2002; Bonner, 2008, p. 161; Ryan & Deci, 2000). Furthermore, employees have a tendency to believe they perform worse than average on difficult tasks, which reduces motivation to exert effort, while they tend to believe they perform better than average on easy tasks, which increases their motivation to exert effort (Moore, 2007). Overall, task complexity reduces employee performance (Bonner & Sprinkle, 2002). However, complex tasks require additional effort and may benefit from undue focus on a single task (Bonner & Sprinkle, 2002). There are multiple ways to solve a complex task and it is unclear which strategy is the best way (Farkas, 2013, p. 25). Extra time dedicated to a single task allows employees to test different strategies in order to find a better strategy. The quality of the strategy used compared to the effort exerted is more important for the successful handling of complex tasks than the handling of simple tasks (Locke & Latham, 1990). If employees' effort allocation is distorted and they overly focus on a single task, this extra effort allocation can be beneficial for solving a complex task as it allows employees to search for a better strategy to deal with the task. As such, the negative effect of distorted allocation of effort across tasks on employee overall performance is less negative for complex tasks than for simple tasks. The reasoning above leads to the following hypotheses:

Hypothesis 1: A distorted allocation of effort across tasks leads to lower employee overall performance on both tasks.

Hypothesis 2: The negative effect of distorted allocation of effort across tasks on overall employee performance on both tasks is less negative in a context of high task complexity than in a context of low task complexity.

Relative performance information, managers' bonus allocation and employees' effort allocation across tasks

Effort-averse employees will allocate effort to the task for which they have higher skills (Farkas, 2013). Employees typically spend more time on easier, more productive tasks in order to maximize the return to effort (Duggan et al., 2013; Janssen et al., 2011; Payne et al., 2007). The reasoning above leads to the following hypotheses:

[59] According to self-determination theory, extrinsically motivated employees perform a task because of some external reward. Intrinsically motivated employees perform a task because it provides them with inherent satisfaction, fun or challenge instead of some external reward (Ryan & Deci, 2000).

Hypothesis 3a: In the first period, employees will focus on the task for which they have the highest skills.

Hypothesis 3b: Employees will focus on the easiest task in order to perform at least well on that task.

However, organizations have a certain preferred effort allocation in mind. Employees' tendency to focus on easier, more productive tasks or tasks for which they have higher skills causes a distortion in effort allocations across tasks away from firm-preferred effort allocations. Organizations then use management control systems (MCS) in order to align employees' goals with organizational goals and strategies (Chenhall, 2003; Merchant & Van der Stede, 2007; Salterio, 2015). MCS measure, analyze and report information that is meant to be useful to managers and employees in order to perform their job well (Otley. 1999; Sprinkle, 2003; Sprinkle & Williamson, 2006). The MCS consists of tools such as performance overviews containing detailed RPI and managers' performance evaluation and bonus allocation that can steer employees towards firms objectives and preferences (Chenhall, 2003). Firms often allow managers to evaluate employees' performance by means of a performance overview containing multiple performance measures, such as a balanced scorecard. The manager then decides upon the weights she attaches to the different performance measures (Cardinaels & van Veen-Dirks, 2010; Höppe & Moers, 2011; Ittner et al., 2003). Therefore subjective performance evaluation differs from formula-based bonus allocations, which are based on weights that are fixed ex-ante. Firms often install a discretionary performance evaluation system in order to enhance organizational performance (Baiman & Rajan, 1995; Franco-Santos et al., 2012; Gibbs et al., 2004). In a multitasking environment, a discretionary system in which the manager has the final say in the evaluation seems particularly useful, since a complete contract specifying all possible outcomes and related employee rewards would be impossible (Gibbs et al., 2004; MacLeod and Parent, 1999; Rajan & Reichelstein, 2006). With subjective evaluations the manager can redirect employees to the performance measures that need more attention in the future due to changing circumstances. Additionally, the manager can avoid that employees manipulate measures or focus too much on particular performance measures (Bol, 2008). Discretionary bonus pools mitigate the occurrence of distorted performance, which is often associated with formula-based performance evaluations (Ittner et al., 2003). Equity theory predicts that managers will take into account employees relative performance when allocating the bonus and as such their evaluation will be in line with the RPI (Adams, 1963, 1965; Walster et al., 1973). Managers will consider firms' objectives and preferences in their bonus allocation as well (Chenhall, 2003). The RPI allows managers to thoroughly analyze employees' performance, to detect and to punish employees' distorted performance and to mitigate dysfunctional behavior in the future (Baker et al., 1994; Ittner et al., 2003). As such, we predict that managers' bonus allocation will consider distorted effort allocations away from firm preferences as well. The reasoning above leads to the following hypotheses:

Hypothesis 4a: Managers' bonus allocation is in line with the RPI they receive about the employees.

Hypothesis 4b: Ceteris paribus, managers' bonus allocation will be lower when employees' performance deviates from the firm-preferred effort allocation.

The provision of RPI by the MCS allows employees to compare their performance with their peers.[60] Humans have an innate desire to compare themselves with others (Festinger, 1954). Social comparison leads to competitiveness, which will eventually affect employee performance and employee effort allocation across tasks (Eichhorn, 2016; Festinger, 1954; Steele, 1988). There are two types of competitiveness arising from RPI: 1) goal, personal development or task-orientated competiveness (the focus of social comparison theory) and 2) hyper-, interpersonal or other-referenced competitiveness (the focus of self-affirmation theory). How competitiveness affects employee performance and employee effort allocation across tasks depends on the specific type of competitiveness that is provoked (task-oriented or other-referenced competitiveness) and their related rivalry theories (social comparison theory or self-affirmation theory) (Eichhorn, 2016; Festinger, 1954; Steele, 1988).

Social comparison theory states that people are willing to outperform others (Festinger, 1954). Employees tend to behave competitively, in the sense that they strive for better overall performance than that of their colleague in order to generate positive feelings and a positive self-image (Festinger, 1954; Hannan et al., 2013). Task-oriented competiveness thereby focuses on self-improvement. This type of competitiveness leads to an increase in task-related effort and effort aimed at learning in order to improve performance (Eichhorn, 2016). According to social comparison theory, social comparison leads to competition and an unidirectional drive upward as employees want to outperform their colleagues by performing the tasks better (Festinger, 1954, P. 124; Tafkov, 2013). Social comparison theory further predicts a trend towards uniformity, a tendency to change one's own position so as to move closer to others in the group when a discrepancy in abilities exists (Eichhorn, 2016; Festinger, 1954, p.126). The provision of detailed RPI indicates employees on which task(s) they outperformed their colleague or whether they overly focused on one task at the expense of the other task (distorted performance). This information allows employees to discover how they should shift their time allocation across tasks in order to allocate their effort in line with firm preferences in order to compete with their colleague for the financial reward and outperform their colleague. Social comparison thereby predicts that employees will shift their time allocation from the task on which they outperformed their colleague (the task on which they overly focused) to the task for which they underperformed their colleague (the task they neglected), in line with firm-preferences in order to outperform their colleagues on all tasks. By doing so employees allocate their effort more balanced across tasks.

However, self-affirmation theory would make very different predictions. Self-affirmation theory focuses on other-referenced competiveness that is concerned with maintaining a positive

[60] This social comparison can lead to assimilation (the belief that one can achieve the same performance as the comparison target) or contrast (the belief that you and the target are very different). Assimilation increases motivation when comparing oneself with a better-performing person, but can cause demotivation when comparing oneself with a worse-performing person. A contrast leads to the opposite effect. It leads to demotivation when comparing oneself with a better-performing person because it leads to negative feelings of inferiority and lack of ability, but it creates positive feelings when comparing oneself with a worse-performing person. Overall, assimilation should dominate as people cease social comparison when they are very different from each other (Eichhorn, 2016).

self-image, outperforming competitors and obtaining a higher rank instead of performing a task as good as possible (Eichhorn, 2016; Steele, 1988). This theory predicts that social comparison causes employees to overly focus on some tasks at the expense of other tasks, because they want to address a threat to their self-image in one area by showing their competence in another area (Steele, 1988) and/or satisfy an innate desire for social distinction (Frey, 2007) by outperforming their peers in at least one area (Farkas, 2013; Hannan et al., 2013). Employees that perform better than others experience positive feelings such as pride, whereas employees that perform worse than others experience negative feelings such as shame (Hannan et al., 2013; Lazarus, 1991; Smith, 2000). Therefore, self-affirmation theory predicts that employees will focus on the tasks for which they have the highest skills in order to secure at least good performance on those tasks and to distinguish themselves from their peers at least at one performance dimension instead of focusing on overall employee performance (Hannan et al., 2013).

Empirical research on the effect of the provision of RPI by the MCS on employee performance is inconclusive. Consistent with social-comparison theory, in a single-task setting RPI leads to better employee performance (Kramer et al., 2016; Murthy & Schafer, 2011; Tafkov, 2013) and this positive effect is larger under an individual performance-based compensation contract than under a flat-wage contract (Tafkov, 2013). Consistent with self-affirmation theory, in a multi-tasking setting, RPI reduces employee performance and it leads to distorted effort allocations (Hannan et al., 2013). However, Hannan et al. (2013) does not consider financial incentives tied to employee performance, but only financial incentives to allocate effort according to firm preferences. Expectancy theory would predict that employees' behavior changes when financial incentives are at stake. With financial incentives linked to their performance, employees will be motivated to exert extra effort and to allocate their effort according to firm preferences in order to increase the likelihood of obtaining a bonus (Vroom, 1964). However, Farkas (2013) demonstrates that financial incentives are not effective at directing effort or in improving performance and RPI has a negative effect on employees' effort allocation and performance as well. This result may be due to the fact that employees did not believe they were able to attain the pre-specified performance goal on the firm-preferred complex task in order to qualify for the goal-based compensation, whereas employees under a discretionary bonus pool do not have to attain a certain performance level before they might qualify for a bonus (Farkas, 2013). Therefore, we still believe that the management control system will motivate the appropriate allocation of effort across tasks according to firm preferences (Chenhall, 2003; Farkas, 2013; Hannan et al., 2013). Employees will assess the likelihood that increased effort will yield higher rewards and will allocate their effort accordingly (Farkas, 2013; Vroom, 1964). This effort directing role of incentives provides a powerful control tool, in that it allows the manager to direct employee effort toward the tasks most crucial for firm success (Bonner & Sprinkle, 2002; Brüggen & Moers, 2007; Hecht et al., 2012; Holmström & Milgrom, 1991). Employees that distort their effort towards a single task, will receive a smaller bonus (see hypothesis 4b). The discretionary bonus thereby limits the opportunity for self-affirmation in a single area of performance and reduces the occurrence of disruptive employee behavior. Therefore we predict that a manager's performance evaluation

will lead employees to allocate their effort across tasks in line with firm preferences and social comparison theory. The reasoning above leads to the following hypotheses:

Hypothesis 5: After the provision of RPI and the performance evaluation, employees reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague.

4.3 Method

We used two computer-based experiments in order to test our hypotheses. Participants were grouped in anonymous groups of three persons in which two employees work for a manager for two periods. Each period the employees have to work individually on two independent but equally important tasks. The tasks consist of solving two numerical puzzles (a binary puzzle and a Japanese puzzle), and the employees' job is to provide as many correct answers as possible to both puzzles. The employees could freely switch between both tasks and they could freely decide how much time to allocate to each task separately. It was specified that both tasks were equally important for the manager, indicating that the firm-preferred allocation of effort is an equal allocation across the two tasks. The total time available in a period to perform both tasks is constrained at ten minutes. Participants were familiarized with the puzzles by means of practice puzzles before the actual experiment took place. Figure 4.1 displays the practice puzzles, with the related rules and solutions for the binary puzzle and Japanese puzzle respectively.

After each period the manager and the employees receive detailed RPI on both tasks for both employees. The manager can then base her performance evaluation on the performance measures productivity (i.e. the number of answers that each employee provided for each task) and accuracy (i.e. the percentage of correct answers that each employee provided for a sample of 5 randomly selected answers for each task). Therefore, the managers' imperfect measure of employees' performance on each task is the product of the variables productivity and accuracy (*"ImperfectPerformance$_{it}$"*). The manager is required to evaluate and reward employee performance on both tasks by means of a bonus pool, and she has the possibility to justify her decision to each employee. The manager does not need to allocate the entire bonus pool.

In the first experiment, managers receive after each period performance information for both employees on both tasks jointly and they evaluate, reward and (possibly) explain their decision for the performance on both tasks jointly (allocating one bonus pool). In the second experiment, managers allocate two separate bonus pools. They first receive after each period performance information for both employees on task 1 and they evaluate, reward and (possibly) explain their evaluation decision for the first task. After their bonus allocation for task 1, they receive performance information for both employees on task 2 and they allocate a second bonus then.

Figure 4.1 Practice puzzles: rules, example and solution

Binary puzzle

Rules:
(1) Each cell contains a “0” or a “1”.
(2) No more than two the same numbers may be placed next to each other or beneath each other.
(3) Each row and each column contains as much “0” values as “1” values.
(4) Each row and each column is unique, but any row may be filled similarly as any column.
(5) Each binary puzzle has a unique solution. This solution can always be found without guessing.

Example

	1				
0				1	1
0	0			1	
		1			0
			1		
1				0	

Solution

1	1	0	1	0	0
0	1	0	0	1	1
0	0	1	0	1	1
1	0	1	1	0	0
0	1	0	1	1	0
1	0	1	0	0	1

Japanese puzzle

Rules:
(1) Each cell contains a “0” or a “1”.
(2) The numbers left to the rows indicate how many adjacent cells contain a “1” in that row, after which at least one cell must contain a “0” e.g. When “1 2” is displayed to the left of a row, then (possibly after one or multiple cells with a “0”) one cell should contain a “1”. Next, at least one cell must contain a “0” and afterwards two adjacent cells should contain a “1”.
(3) The numbers above each column indicate how many adjacent cells contain a “1” in that column, after which at least one cell must contain a “0” e.g. When “3 2 1” is displayed above a column, then (possibly after one or multiple cells with a “0”) first three adjacent cells should contain a “1”, next, at least one cell must contain a “0” and afterwards two adjacent cells should contain a “1”. Next, again at least one cell must contain a “0”, after which one cell must contain a “1”.
(4) Each Japanese puzzle has a unique solution. This solution can always be found without guessing.

Example

	3 1	1 1 1	1 3		5
3 1					
1 1					
3 1					
1 1					
3 1					

Solution
(black cells = “1”, white cells = “0”)

	3 1	1 1 1	1 3		5
3 1	■	■	■		■
1 1	■				■
3 1	■	■	■		■
1 1			■		■
3 1	■	■	■		■

In all experimental conditions the total bonus that could be allocated by a manager is the same, 20.000 points. In the first experiment ('one bonus pool' experiment) the bonus that could be allocated consists of 20.000 points, while in the second experiment ('two bonus pools' experiment) each bonus that could be allocated consists of 10.000 points.[61] The manager's goal is to let her two employees provide as many correct answers as possible to the puzzles and both tasks are equally important. The employees' goal is to receive as many bonus points as possible through the manager's performance evaluations. We conducted the experiment via the Ztree experimental software (Fischbacher, 2007). In both experiments, we manipulate 'task complexity' as explained in the experimental design section.

Experimental design

We conducted two computer-based experiments in which we randomly assigned student participants to a group of three persons in one of the four experimental conditions. The experiments were anonymous, so participants did not know their group members. In both experiments, we manipulated the complexity of the two tasks (the binary and Japanese puzzle) performed by the employees between groups (*Task Complexity*; low vs. high). Both employees that work for the same manager receive the same level of task complexity (either low complexity or high complexity). In the low (high) complexity conditions participants solve a medium-difficult (difficult) binary puzzle and a medium-difficult (difficult) Japanese puzzle[62]. The type of tasks performed in the high and low complexity conditions is the same (task 1 (2) consists of solving a binary (Japanese) puzzles), but we change the complexity of the tasks.

Experimental procedures

Upon arrival at the computer room each participant was randomly assigned to one of the two experimental conditions of the two experiments. Participants were grouped by 3 and received computerized instructions on the experimental task. Two persons assumed the role of employee in a department and one person assumed the role of department manager. The instructions specified that the two employees worked individually on the same two tasks (a binary puzzle and a Japanese puzzle) for the manager during two periods of ten minutes. After the first period, employees received new puzzles to solve. Employees were able to choose how much time they would dedicate to each task and when and how often they would switch between tasks. It was specified that both tasks were equally important to the manager (firm-preferred effort allocation) and that the employees had to provide as many correct answers as

[61] We conducted these two experiments to see whether our results hold in different MCS environments. The bonus system in the first experiment rewards overall performance on both tasks and strong performance on one task can compensate for weak performance on another task. It allows employees to outrank their colleague at least at one task without necessarily performing well on both tasks and should therefore stimulate other-referenced competitiveness (the focus of self-affirmation theory). The bonus system in the second experiment rewards performance on each individual task and weak performance on one task will not (or to a smaller extent) be compensated by strong performance on another task. It should therefore stimulate task-oriented competitiveness (the focus of social comparison theory). This is similar to the different effect of unidimensional and multidimensional RPI on employee performance in Eichhorn (2016).

[62] Based on the indicated difficulty levels on the websites, we selected medium-difficult and difficult binary puzzles (http://www.binairepuzzel.net) and Japanese puzzles (http://pic-a-pix.nl). Additionally, we pilot tested the difficulty levels of the puzzles involving colleagues at the university.

possible. Participants were informed that the more correct answers the employees would provide, the better the department would perform and the higher the chances were for the manager to win a 15 euros reward. Furthermore, it was specified that after each period, the manager would evaluate the performance of each employee and that she was able to allocate one bonus pool of 20.000 points (experiment 1) or two bonus pools of 10.000 points each (experiment 2) amongst the employees. The manager had the discretion to decide whether and how many bonus points a particular employee would receive. The more bonus points an employee received, the higher the chances were for that employee to win a 15 euros reward, a lottery as in Masschelein et al. (2012). It was specified that the manager did not have to divide all bonus points, but non-allocated bonus points would not benefit the manager, nor the employees. This is in line with for example Bol et al. (2016), where managers did not have to allocate the entire bonus either. This set-up helps avoiding agency costs related to predefined bonus pools, such as collusion and employee opportunism (Fisher et al., 2005). An employee did only know how many bonus points he received and the total amount of bonus points that could be divided. He would not know how many bonus points his colleague received. This is in line with pay secrecy practices in many firms in the US and abroad (Bamberger & Belogolovsky, 2010; Belogolovsky & Bamberger, 2014; Colella et al., 2007; Day, 2007; Thompson & Pronsky, 1975). Additionally, the manager could provide a written justification of her bonus allocation to each employee separately. Each employee could only see the written justification addressed to him, not the written justification addressed to his colleague.

The instructions further indicated that all participants received a performance overview indicating how many answers an employee provided; this was per period for each employee and for both tasks. Furthermore, a random sample of five answers per task from each employee was used to calculate the percentage of correctly answered numbers for that sample. The manager received this performance overview to use for her bonus allocation. In the first experiment (‘one bonus pool’) the manager could allocate and justify a 20.000 points bonus amongst the two employees based on the performance information for the binary puzzle and the Japanese puzzle. In the second experiment (‘two bonus pools’) the manager could first allocate and justify a 10.000 points bonus amongst the two employees based on the performance information for the binary puzzle. Next, she could allocate and justify a 10.000 points bonus amongst the two employees based on the performance information for the Japanese puzzle.

After reading all case information, all participants were given two practice puzzles and the rules to solve those puzzles (see Figure 4.1) in order to familiarize with the required experimental task. Participants received a paper printed version of the rules as well. Only after all three participants in a group had completed the practice puzzles, the first period of the actual experiment started. Participants needed on average 235 (302) seconds to familiarize with the rules and the practice puzzles of the binary (Japanese) puzzle. Next, the experiment started. After 10 minutes the manager allocated the bonus and possibly justified her bonus allocation decision, after which the second period started. In the end, participants received a number of post-experimental and demographic questions.

Variables

Employee performance, employee effort allocation and managers' bonus allocation

The manager's goal is to have her two employees provide as many correct answers as possible to the binary and Japanese puzzle, both tasks being equally important. The variable $EmployeePerformance_{it}$ captures employees' overall performance. It measures the actual number of correct answers both employees provided each period for task 1 (the binary puzzle) and task 2 (the Japanese puzzle) together. However, the actual employee performance is unknown to the manager and the employees as the performance overview is based on a sample which only imperfectly captures employees' performance. The performance overview contains the performance measures productivity (i.e. the number of answers that each employee provided for each task) and accuracy (i.e. the percentage of correct answers that each employee provided for a sample of 5 randomly selected answers for each task). Therefore, the managers' imperfect measure of employees' performance on each task is the product of the variables productivity and accuracy (*"*$ImperfectPerformance_{it}$*"*).

We further define the variables $RelativePerfTask1_{it}$ and $RelativePerfTask2_{it}$. They indicate how an employee performed relative to his colleague on task 1 or task 2. Positive values signal an employee outperformed his colleague on task 1 or task 2. $RelativePerfTask1_{it}$ is calculated as ($ImperfectPerformance_{it}$ on task 1 for an employee minus $ImperfectPerformance_{it}$ on task 1 for his colleague) divided by the sum of $ImperfectPerformance_{it}$ on both tasks for both employees. $RelativePerfTask2_{it}$ is calculated as ($ImperfectPerformance_{it}$ on task 2 for an employee minus $ImperfectPerformance_{it}$ on task 2 for his colleague) divided by the sum of $ImperfectPerformance_{it}$ on both tasks for both employees.

Each period employees were free to allocate 10 minutes (600 seconds) across the two puzzles. $TimeSpentTask1_{it}$ equals the number of seconds an employee spent on task 1. The number of seconds an employee spent on task 2 is then 600 minus $TimeSpentTask1_{it}$. This measure captures employees' allocation of effort across tasks. $\Delta\ TimeSpentTask1_{i}$ is the change in the number of seconds an employee spent on task 1 from the first period to the second period. Positive values signal an increase in the time spent on task 1.

After each period the manager was then able to allocate a bonus of 20.000 experimental units (experiment 1) or two bonuses of each 10.000 experimental units (experiment 2) amongst the two employees. $Bonus_{it}$ measures the number of experimental units each employee received each period.

Task complexity and distorted effort allocation

Complexity is a dummy variable related to the manipulation of task complexity in the two experiments. A value of zero for this measure indicates that the participants solved medium-difficult puzzles. A value of one on this measure indicates that the participants solved difficult puzzles.

Furthermore, the experimental description indicated that both tasks (binary or Japanese puzzle) were equally important to the manager. We define three variables that measure the extent to which employees' effort allocations across tasks differ from the firm-preferred equal allocation of effort across tasks: $EffortDistortion_{it}$, $PerformanceDistortion_{it}$ and $DistortionImperfectPerformance_{it}$. First, $EffortDistortion_{it}$ is measured as the absolute value of

the difference between the time spent on a single task and the average time available to spend on each task. Second, $PerformanceDistortion_{it}$ is an alternative variable that measures how much employees' performance across tasks deviates from firm preferences for an equal performance across tasks. For each task (the binary and Japanese puzzle respectively) we divide the number of answers an employee provided by the maximum number of answers that could be provided for that task. This leads for each task to a percentage which indicates the extent of completion of that task. $PerformanceDistortion_{it}$ is then the absolute value of the difference between the percentages for both tasks. Third, $DistortionImperfectPerformance_{it}$ is the absolute value of the difference between $ImperfectPerformance_{it}$ on task 1 and $ImperfectPerformance_{it}$ on task 2. It measures the degree of performance distortion from an equal allocation of performance across both tasks, as perceived by the manager based on the RPI. Finally, $Complexity * EffortDistortion_{it}$, and $Complexity * PerformanceDistortion_{it}$ are the interactions between *Complexity* and $EffortDistortion_{it}$ ($PerformanceDistortion_{it}$) respectively and $ImperfectPerformance * DistortionImperfectPerformance_{it}$ is the interaction of the variables $ImperfectPerformance_{it}$ and $DistortionImperfectPerformance_{it}$.

Control variables

$PracticeTask1_i$ refers to how much time (in seconds) it took the employee to solve the practice puzzle for task 1 (the binary puzzle). $PracticeTask2_i$ refers to how much time (in seconds) it took the employee to solve the practice puzzle for task 2 (the Japanese puzzle). Furthermore, we asked employees in the post-experimental questionnaire (using a seven-point Likert scale) to indicate the extent to which employees perceived task 1 (the binary puzzle) difficult ($DifficultyTask1_i$) and the extent to which they perceived task 2 (the Japanese puzzle) difficult ($DifficultyTask2_i$). A high value for $PracticeTask1_i$, $PracticeTask2_i$, $DifficultyTask1_i$ or $DifficultyTask2_i$ indicates less ability for that particular task.

Participants

We recruited 348 undergraduate students from a business program in a large university. The students received a course credit as an incentive to participate in the experiments. Furthermore, 20 managers and 40 employees received a 15 euros reward based on a lottery with increasing chances to win depending on their performance during the experiments. The experiments lasted 50 to 90 minutes (depending on the decision speed of the participants). 166 students were female and 182 were male. They were 22 years old on average (minimum 19 years, maximum 30 years) and they had 10.1 months of work experience, which was significantly different from 0 ($p < 0.01$).

4.4 Results

Manipulation checks

In order to check the manipulation of *Task Complexity*, we asked participants whether they experienced the puzzles that the employees had to solve as relatively easy or relatively difficult. 94% of the participants in the low complexity conditions indicated that the puzzles were relatively easy, whereas 99% of the participants in the high complexity conditions indicated

that the puzzles were relatively difficult. Additionally, we asked participants to indicate (on a seven-point scale) their agreement with the statements that task 1 (the binary puzzle) was difficult and that task 2 (the Japanese puzzle) was difficult, respectively. Participants in the high complexity conditions agreed significantly more with the statements than participants in the low complexity conditions ($F_{1,\ 346}$ = 153.04; p <0.01 and $F_{1,\ 346}$ = 36.42; p <0.01 respectively).[63] Lastly, we asked participants who took the role of an employee to indicate (on a seven-point scale) whether solving the puzzles was a difficult task to execute (on a seven-point scale). Employees judged the task difficulty to be significantly higher in the high complexity conditions than in the low complexity conditions ($F_{1,\ 230}$ = 93.003; p <0.01). We therefore conclude that the manipulation of T*ask Complexity* was successful.

Additionally we asked participants to indicate (on a seven-point scale) their agreement with a couple of statements related to the experimental task. Participants indicated they were motivated. The mean of this measure is 6.17 on 7, which is significantly higher than the scale midpoint of 4 (t_{347} = 39.46; p <0.01). The participants said that the instructions were clear. The mean of this measure was 5.70 on 7, which is significantly higher than the scale midpoint of 4 (t_{347} = 25.95; p <0.01). Participants indicated that they had enough time to make their decisions (mean of 4.85 on 7, which is significantly higher than the scale midpoint of 4 (t_{347} = 9.80; p <0.01)). Together, these results suggest that participants felt they had sufficient time and were capable of making a reasonable performance evaluation.

Descriptive statistics

Table 4.1 provides the definitions and descriptive statistics mean (standard deviation) of our main dependent variables used to test our hypotheses by task complexity for the two experiments. The table clearly indicates that the manipulation of task complexity worked as the number of correct answers provided by the employees is much lower in the high complexity conditions than in the low complexity conditions. Further, we can see that on average employees preferred to work on task 1 (the binary puzzle) over task 2 (the Japanese puzzle). Each period, employees could freely decide how to allocate their time (600 seconds per period) across the two tasks. In all experimental conditions employees spent on average more seconds on task 1 than on task 2. Next, we can see that managers did not allocate the entire bonus. Managers in all experimental conditions could divide a bonus of 20.000 experimental points across the two employees, which means an average of 10.000 points per employee. However, the average allocated bonus ranges between 7,572.54 and 8,640.05. Finally, we can see that, on average, employees spent less time on task 1 in the second period than in the first period.

[63] Every p-value mentioned in this paper is a two-sided p-value.

Table 4.1 Descriptive Statistics

Mean (Standard Deviation)

	Experiment 1		Experiment 2	
	Task Complexity[a]		Task Complexity[a]	
	Low	High	Low	High
Dependent Measure	(n = 58)[b]	(n = 56)	(n = 60)	(n = 58)
EmployeePerformance[c]	119.45	81.14	128.78	81.03
	(36.25)	(28.23)	(34.67)	(27.34)
TimeSpentTask1[d]	358.18	315.89	365.58	327.25
	(111.91)	(146.79)	(93.13)	(132.77)
Bonus[e]	8,640.05	7,572.54	8,276.13	7,857.70
	(3,439.12)	(3,886.69)	(3,750.96)	(3,343.27)
Δ TimeSpentTask1[f]	-69.44	-103.18	-35.80	-90.85
	(138.73)	(151.56)	(100.84)	(123.42)

[a] *Task Complexity* refers to whether the complexity of the puzzles is high or low. Task Complexity is coded as 1 for the High Complexity conditions and 0 for the Low Complexity conditions.
[b] N refers to the number of employees in each condition.
[c] *EmployeePerformance* is the average number of correct answers an employee provided to both puzzles over the two periods.
[d] *TimeSpentTask1* equals the average number of seconds an employee spent on task 1 over the two periods. *TimeSpentTask2* is then 600 seconds minus *TimeSpentTask1*.
[e] *Bonus* is the average number of experimental units an employee received as a bonus from his manager over the two periods.
[f] *Δ TimeSpentTask1* is the change in the average number of seconds an employee spent on task 1 from the first period to the second period. Positive values signal an increase in the time spent on task 1.

Hypothesis tests

Hypothesis 1 states that a distorted allocation of effort across tasks leads to lower employee overall performance on both tasks and hypothesis 2 predicts that the negative effect of a distorted allocation of effort across tasks on overall employee performance on both tasks is less negative for complex tasks than for simple tasks. In order to test hypotheses 1 and 2, we conducted random-effects regressions for both experiments with $EmployeePerformance_{it}$ as the dependent variable and *Complexity*, $EffortDistortion_{it}$ and their interaction as the independent variables (see Panel A of Table 4.2). We further controlled for the period and employees' abilities for each task (*PracticeTask1* and *PracticeTask2*) and we clustered the standard errors by the manager. Hypothesis 1 indicates we expect a negative effect of *EffortDistortion* on *EmployeePerformance* and hypothesis 2 indicates we expect a positive effect of the interaction of *Complexity* and *EffortDistortion* on *EmployeePerformance*. Additionally we expect a negative effect of *PracticeTask1* and *PracticeTask2* on *EmployeePerformance*, as a lack of skills should reduce performance.

As reported in Panel A of Table 4.2, the results for both experiments provide clear support for H1, as *EffortDistortion* is negatively related to *EmployeePerformance* (Experiment 1: t = -10.37; p<0.001; Experiment 2: t = -11.63; p<0.001). The results further provide support for H2, as the interaction between *Complexity* and *EffortDistortion* is positively related to *EmployeePerformance* (Experiment 1: t = 3.90; p<0.001; Experiment 2: t = 4.62; p<0.001).

Additionally, we find a negative effect of *PracticeTask1* (Experiment 1: t = -2.53; p=0.011; Experiment 2: t = -3.58; p<0.001) and *PracticeTask2* (Experiment 1: t = -5.10; p<0.001; Experiment 2: t = -3.40; p=0.001) on *EmployeePerformance* indicating that employees that have less ability performed worse. Finally, we find a positive effect of *Period* on *EmployeePerformance* (Experiment 1: t = 5.15; p<0.001; Experiment 2: t = 5.60; p=0.001), indicating learning effects, strategy development and improved performance that was amongst others predicted by social comparison theory (Bonner & Sprinkle, 2002; Festinger, 1954; Locke & Latham, 1991). In Panel B of Table 4.2 we further did a robustness check for hypotheses 1 and 2 by conducting random-effects regressions for both experiments with $EmployeePerformance_{it}$ as the dependent variable and *Complexity*, $PerformanceDistortion_{it}$ and their interaction as the independent variables (see Panel B of Table 4.2). Our results remain similar.

Hypothesis 3a predicted that, in the first period, employees will focus on the task for which they have the highest skills. Hypothesis 3b further predicted that Employees will focus on the easiest task in order to perform at least well on that task.

In order to test hypothesis 3a, we conducted OLS-regressions for both experiments with $TimeSpentTask1_i$ observed in the first period as the dependent variable and employees' abilities for each task (*PracticeTask1* and *PracticeTask2*) as the independent variables (see Panel A of Table 4.3). The employees' abilities for each task jointly affect the time spent on the first task in the first period (Experiment 1: $F_{2, 111} = 5.27$; p = 0.007; Experiment 2: $F_{2, 115} = 9.84$; p <0.001) and this result is mainly driven by employees' ability for task 2 (*PracticeTask2*). The less ability employees have for task 2 the Japanese puzzle (i.e. the more time the employees needed to solve the practice puzzle for task 2 (the Japanese puzzle)), the more time they spend on the other task (the binary puzzle) in the first period (Experiment 1: t = 2.53; p=0.013; Experiment 2: t = 4.18; p<0.001). Employees ability for task 1 (the binary puzzle) (*PracticeTask1*) does not affect their focus on the binary puzzle in the first period (Experiment 1: t = 1.09; p=0.278; Experiment 2: t = -0.28; p=0.780). These results provide support for hypothesis 3a.

As a robustness test for hypothesis 3a we further conducted OLS-regressions for both experiments with $TimeSpentTask1_i$ observed in the first period as the dependent variable and employees' perceptions of the difficulty of each task (*DifficultyTask1* and *DifficultyTask2*) as the independent variables (see Panel B of Table 4.3). The employees' perceptions of task difficulty jointly affect the time spent on the first task in the first period (Experiment 1: $F_{2, 111} = 7.35$; p = 0.001; Experiment 2: $F_{2, 115} = 7.32$; p = 0.001). The more difficult employees perceived task 1, the less time they spent on task 1 in the first period (Experiment 1: t = -2.57; p=0.011; Experiment 2: t = -2.15; p=0.034) and the more difficult employees perceived task 2, the more time they spent on task 1 in the first period (Experiment 1: t = 3.37; p=0.001; Experiment 2: t = 3.53; p=0.001), which supports hypothesis 3a as well.

In order to test hypothesis 3b, we asked employees in the post-experimental questionnaire **(**using a seven point Likert scale), to indicate whether they focused on the easiest task in order to perform at least well on that task (see Panel C of Table 4.3). In line with hypothesis 3b, participants agreed with this statement. The mean response (4.56 (4.53) on 7 for experiment 1 (experiment 2)) is significantly higher than the scale midpoint of 4 ($t_{113} = 2.94$; p <0.01; ($t_{117} = 2.83$; p <0.01)). We therefore provide support for H3b.

Hypothesis 4a and 4b claimed that when managers allocate their bonus, they take into account the RPI and the firm-preferred equal allocation of effort. Hypothesis 4a predicts that the bonus allocated to the employee is positively related to employees' relative performance and hypothesis 4b predicts that the bonus allocated to the employee is negatively related to the extent to which employee performance deviates from the firm-preferred effort allocation. When evaluating performance, the true employee performance on both tasks is unknown to the manager. The manager can only consult the performance measures productivity (i.e. the number of answers that each employee provided for each task) and accuracy (i.e. the percentage of correct answers that each employee provided for a sample of 5 randomly selected answers for each task). Therefore, the managers' imperfect measure of employees' performance on each task is the product of the variables productivity and accuracy (*"ImperfectPerformance$_{it}$"*). In order to test hypotheses 4a, we conducted random-effects regressions for both experiments with *Bonus$_{it}$* as the dependent variable and *RelativePerfTask1$_{it}$* and *RelativePerfTask2$_{it}$* as the independent variables (see Panel A of Table 4.4). We further controlled for the period.

As reported in Panel A of Table 4.4, the results for both experiments provide clear support for H4a, as both *RelativePerfTask1* (Experiment 1: $t = 2.97$; $p = 0.003$; Experiment 2: $t = 4.18$; $p<0.001$) and *RelativePerfTask2* (Experiment 1: $t = 6.10$; $p<0.001$; Experiment 2: $t = 7.14$; $p<0.001$) are positively related to *Bonus*. Managers consider employees relative performance on both tasks when allocating their bonus.

In order to further test hypotheses 4a and in order to test 4b, we conducted random-effects regressions for both experiments with *Bonus$_{it}$* as the dependent variable and *ImperfectPerformance$_{it}$*, *DistortionImperfectPerformance$_{it}$* and their interaction as the independent variables (see Panel B of Table 4.4). In line with hypothesis 4a, we find that the bonus the managers allocate is positively related to employees performance (*ImperfectPerformance*) (Experiment 1: $t = 7.00$; $p<0.001$; Experiment 2: $t = 8.45$; $p<0.001$). However, in contrast to our hypothesis 4b, managers do not consider firm preferences for an equal allocation of effort/performance. *DistortionImperfectPerformance* (Experiment 1: $t = 1.10$; $p = 0.271$; Experiment 2: $t = -0.23$; $p = 0.818$) and the interaction between *ImperfectPerformance* and *DistortionImperfectPerformance* (Experiment 1: $t = 1.54$; $p = 0.124$; Experiment 2: $t = -0.09$; $p = 0.928$) do not significantly affect managers' bonus. As such, we do find support for hypothesis 4a, but we do not find support for hypothesis 4b.

Lastly, hypothesis 5 predicted that employees would reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague. This reallocation will eventually result in a less distorted allocation of effort across tasks and accordingly this leads to better employee overall performance (see H1 and Table 4.2). In order to test hypothesis 5 we conducted an OLS-regression with *Δ TimeSpentTask1* as the dependent variable and *RelativePerfTask1* and *RelativePerfTask2* in the first period as the independent variables (see Table 4.5). We expect a negative coefficient for *RelativePerfTask1* and a positive coefficient for *RelativePerfTask2*. In line with hypothesis 5, we find that the difference in time spent on task 1 is positively related to employees relative performance on task 2 (*RelativePerfTask2*) (Experiment 1: $t = 2.30$; $p = 0.025$; Experiment 2: $t = 3.47$; $p = 0.001$) and that the difference in time spent on task 1 is negatively related to employees relative performance on task 1 (*RelativePerfTask1*) for experiment 1 ($t = -3.57$; $p = 0.001$), but

surprisingly not for experiment 2 (t = 0.64; p = 0.526). We therefore find (partial) support for hypothesis 5 in experiment 1 (experiment 2).

Table 4.2 Random-effects regression results: H1 & H2

Panel A: H1 & 2: Effect of Effort Distortion and Complexity on Employee Performance

	Experiment 1: Employee Performance[a]			Experiment 2: Employee Performance[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	145.37	23.01	<0.001	146.80	24.30	<0.001
Period[b]	13.18	5.15	<0.001	14.09	5.60	<0.001
Complexity[c]	-49.30	-8.03	<0.001	-55.16	-13.40	<0.001
EffortDistortion[d]	-0.20	-10.37	<0.001	-0.24	-11.63	<0.001
Complexity * EffortDistortion[e]	0.10	3.90	<0.001	0.13	4.62	<0.001
PracticeTask1[f]	-0.02	-2.53	0.011	-0.04	-3.58	<0.001
PracticeTask2[g]	-0.06	-5.10	<0.001	-0.03	-3.40	0.001
# of observations	228			236		
Overall R^2	0.597			0.678		
Prob>chi2	0.000			0.000		

Panel B: H1 & H2: Effect of Performance Distortion and Complexity on Employee Performance

	Experiment 1: Employee Performance[a]			Experiment 2: Employee Performance[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	144.41	20.83	<0.001	147.27	21.98	<0.001
Period[b]	13.45	5.24	<0.001	12.43	5.06	<0.001
Complexity[c]	-53.60	-7.85	<0.001	-65.40	-13.46	<0.001
PerformanceDistortion[h]	-63.70	-8.94	<0.001	-74.55	-11.02	<0.001
Complexity * PerformanceDistortion[i]	45.28	4.51	<0.001	77.65	7.98	<0.001
PracticeTask1[f]	-0.02	-3.08	0.002	-0.04	-2.49	0.013
PracticeTask2[g]	-0.05	-4.89	<0.001	-0.04	-3.52	<0.001
# of observations	228			236		
Overall R^2	0.606			0.657		
Prob>chi2	0.000			0.000		

We used linear random-effects models and we adjusted the standard errors of the estimates for clustering by group of two employees and their manager.

[a] The dependent variable is *Employee Performance*$_{it}$, which is the number of correct answers an employee provided to both puzzles in period t.

[b] *Period* refers to the period in the experiment (1 or 2).

[c] *Complexity* refers to whether the complexity of the tasks the employees performed was high or low.

[d] *EffortDistortion*$_{it}$ equals the absolute value of the difference between the time spent on a single task and the average time available to spend on each task). It measures the degree of effort distortion from an equal allocation of effort across both tasks.

[e] *Complexity* * *EffortDistortion*$_{it}$ refers to the interaction of the variables *Complexity* and *EffortDistortion*$_{it}$.

[f] *PracticeTask1*$_i$ equals the number of seconds it took the employee to solve the practice puzzle for task 1. High values for this variable, indicate less skill for the task.

[g] *PracticeTask2*$_i$ equals the number of seconds it took the employee to solve the practice puzzle for task 2. High values for this variable, indicate less skill for the task.

[h] *PerformanceDistortion*$_{it}$ equals the absolute value of the difference in the percentage of completion of task 1 and task 2. It measures the degree of performance distortion from an equal allocation of performance across both tasks.

[i] Complexity * *PerformanceDistortion*$_{it}$ refers to the interaction of the variables *Complexity* and *PerformanceDistortion*$_{it}$.

Table 4.3 OLS regression and t-test results: H3a & H3b

Panel A: H3a: Effect of Task Ability on Time Spent on Task 1 in the first period

	Experiment 1: TimeSpentTask1[a]			Experiment 2: TimeSpentTask1[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	300.76	10.90	<0.001	305.72	14.03	<0.001
PracticeTask1[b]	0.07	1.09	0.278	-0.02	-0.28	0.780
PracticeTask2[c]	0.19	2.53	0.013	0.28	4.18	<0.001
# of observations	114			118		
Adjusted R^2	0.070			0.131		
F	5.27			9.84		
Prob>F	0.007			<0.001		

Panel B: H3a: Effect of Task Difficulty on Time Spent on Task 1 in the first period

	Experiment 1: TimeSpentTask1[a]			Experiment 2: TimeSpentTask1[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	338.66	8.58	<0.001	325.70	9.38	<0.001
DifficultyTask1[d]	-18.16	-2.57	0.011	-13.04	-2.15	0.034
DifficultyTask2[e]	23.34	3.37	0.001	22.20	3.53	0.001
# of observations	114			118		
Adjusted R^2	0.101			0.098		
F	7.35			7.32		
Prob>F	0.001			0.001		

We used OLS regressions.

[a] *TimeSpentTask1* equals the number of seconds employees spent on task 1 in the first period. *TimeSpentTask2* is then 600 seconds minus *TimeSpentTask1*.

[b] *PracticeTask1* equals the number of seconds it took the employee to solve the practice puzzle for task 1. High values for this variable, indicate less skill for the task.

[c] *PracticeTask2* equals the number of seconds it took the employee to solve the practice puzzle for task 2. High values for this variable, indicate less skill for the task.

[d] *DifficultyTask1* refers to the extent to which employees perceived task 1 (the binary puzzle) difficult (using a seven-point Likert scale).

[e] *DifficultyTask2* refers to the extent to which employees perceived task 2 (the Japanese puzzle) difficult (using a seven-point Likert scale).

Panel C: H3b: t-test whether employees focused on the easiest task in order to perform at least well on that task.

	Mean	Std. Dev.	t-statistic	Df	p-value
Experiment 1	4.56	2.04	2.94	113	0.004
Experiment 2	4.53	2.02	2.83	117	0.006

We used a one-sample t-tests; H0: mean = 4; Ha mean ≠ 4. Employees answered the question (using a seven point Likert scale) whether they focused on the easiest task in order to perform at least well on that task.

Table 4.4 Random-effects regression results: H4a & H4b

Panel A: H4a: The effect of RPI on Managers' Bonus Allocation

	Experiment 1: Bonus[a]			Experiment 2:Bonus[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	7,610.13	14.56	<0.001	7,876.52	20.68	<0.001
Period[b]	337.02	1.24	0.216	129.30	0.75	0.455
RelativePerfTask1[c]	4,754.07	2.97	0.003	6,597.37	4.18	<0.001
RelativePerfTask2[d]	7,121.23	6.10	<0.001	5,935.10	7.14	<0.001
# of observations	228			236		
Overall R^2	0.22			0.29		
Prob>chi2	0.000			0.000		

Panel B: H4b: Effect of Performance Distortion on Managers' Bonus Allocation

	Experiment 1: Bonus[a]			Experiment 2:Bonus[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	8,990.01	17.13	<0.001	8,985.99	23.72	<0.001
Period[b]	-598.69	-2.03	0.042	-610.80	-2.97	0.003
ImperfectPerformance[e]	1,831.35	7.00	<0.001	1,686.64	8.45	<0.001
DistortionImperfectPerformance[f]	267.00	1.10	0.271	-34.29	-0.23	0.818
ImperfectPerformance * DistortionImperfectPerformance[g]	467.19	1.54	0.124	-19.68	-0.09	0.928
# of observations	228			236		
Overall R^2	0.21			0.25		
Prob>chi2	0.000			0.000		

We used linear random-effects models and we adjusted the standard errors of the estimates for clustering by group of two employees and their manager.

[a] The dependent variable is $Bonus_{it}$, which is the number of experimental units an employee received as a bonus from his manager in period t.

[b] *Period* refers to the period in the experiment (1 or 2).

[c] $RelativePerfTask1_{it}$ refers to how an employee performed relative to his colleague on task 1. It is calculated as ($ImperfectPerformance_{it}$ on task 1 for an employee minus $ImperfectPerformance_{it}$ on task 1 for his colleague) divided by the sum of $ImperfectPerformance_{it}$ on both tasks for both employees. Positive values signal an employee outperformed his colleague on task 1.

[d] $RelativePerfTask2_{it}$ refers to how an employee performed relative to his colleague on task 2. It is calculated as ($ImperfectPerformance_{it}$ on task 2 for an employee minus $ImperfectPerformance_{it}$ on task 2 for his colleague) divided by the sum of $ImperfectPerformance_{it}$ on both tasks for both employees. Positive values signal an employee outperformed his colleague on task 2.

[e] $ImperfectPerformance_{it}$ is the standardized $ImperfectPerformance_{it}$ on both tasks for an employee.

[f] $DistortionImperfectPerformance_{it}$ is the standardized absolute value of the difference between $ImperfectPerformance_{it}$ on task 1 and $ImperfectPerformance_{it}$ on task 2. It measures the degree of performance distortion from an equal allocation of performance across both tasks, as perceived by the manager based on the RPI.

[g] $ImperfectPerformance * DistortionImperfectPerformance_{it}$ refers to the interaction of the variables $ImperfectPerformance_{it}$ and $DistortionImperfectPerformance_{it}$.

Table 4.5 OLS regression results: H5

H5: The effect of RPI on Change in Time Spent on Task 1

	Experiment 1: Δ TimeSpentTask1[a]			Experiment 2: Δ TimeSpentTask1[a]		
Independent Variable	Estimate	t-statistic	p-value	Estimate	t-statistic	p-value
Intercept	-86.01	-6.12	<0.001	-62.85	-5.49	<0.001
RelativePerfTask1[b]	-247.36	-3.57	0.001	52.94	0.64	0.526
RelativePerfTask2[c]	103.49	2.30	0.025	135.79	3.47	0.001
# of observations	114			118		
R^2	0.15			0.08		
F	11.62			6.71		
Prob>F	<0.001			0.002		

We used OLS regressions and we adjusted the standard errors of the estimates for clustering by group of two employees and their manager.

[a] The dependent variable is Δ TimeSpentTask1, which is the difference in how much time (in seconds) employees spent on task 1 from the first period to the second period. Positive values signal an increase in the time spent on task 1.

[b] *RelativePerfTask1* refers to how an employee performed relative to his colleague on task 1 in the first period. It is calculated as (*ImperfectPerformance* on task 1 for an employee in the first period minus *ImperfectPerformance* on task 1 for his colleague in the first period) divided by the sum of *ImperfectPerformance* on both tasks for both employees in the first period. Positive values signal an employee outperformed his colleague on task 1 in the first period.

[c] *RelativePerfTask2* refers to how an employee performed relative to his colleague on task 2 in the first period. It is calculated as (*ImperfectPerformance* on task 2 for an employee in the first period minus *ImperfectPerformance* on task 2 for his colleague in the first period) divided by the sum of *ImperfectPerformance* on both tasks for both employees in the first period. Positive values signal an employee outperformed his colleague on task 2 in the first period.

4.5 Conclusion and discussion

This paper examines the joint impact of employees' effort allocation across tasks and task complexity on employees' overall performance in a multitasking environment. Additionally, we research how managers' allocation of a discretionary bonus based on detailed relative performance information (RPI) influences employees' effort allocations across tasks.

Employees are often required to take on multiple tasks and responsibilities, while their time is limited. If employees allocate time to a certain task, this time cannot be used for another task (Eichhorn, 2016; Farkas, 2013). Therefore, time management is an important part of employees' job (Adler & Benbunan-Fich, 2012; Farkas, 2013; Gonzalez & Mark, 2004; Kushleyeva et al., 2005; Payne et al., 2007). The way employees allocate their time across their various tasks and responsibilities affects their performance. If employees overly focus on a single task this impacts their overall performance negatively (Hannan et al., 2013). Furthermore, if employees' effort allocation across tasks deviates from what the firm values based on its strategy, objectives or just the direct revenues related to each task, this may impact firm's performance negatively as well (Farkas, 2013). Organizations can therefore assist their employees by designing management control systems that motivate the appropriate level of

effort and allocation of effort across tasks according to firm preferences (Chenhall, 2003; Farkas, 2013; Hannan et al., 2013). Firms often disperse relative performance information (RPI) amongst their employees to motivate employees. This RPI elicits social comparison and competitiveness, which should affect employee performance positively (Festinger, 1954). In a single-task context, prior research demonstrated the positive effect of RPI on employee performance (Kramer et al., 2016; Murthy & Schafer, 2011; Tafkov, 2013). However, in a multi-task context, the provision of RPI causes employees to distort their effort allocation across tasks away from firm preferences in an attempt to socially distinct themselves from their peers (Frey, 2007) by outperforming their peers in at least one area (Farkas, 2013; Hannan et al., 2013). This behavior is in line with self-affirmation theory (Steele, 1988).

In this study we focus on a management control system consisting of a discretionary bonus pool that accompanies the provision of detailed RPI in order to redirect employees' effort towards firm-preferred effort allocations across tasks. Consistent with expectancy theory and social comparison theory (Festinger, 1954; Vroom, 1964), we predict that the detailed RPI and the allocation of the discretionary bonus guide employees towards an allocation of effort across tasks in line with firm preferences.

We investigate our research questions by means of two experiments in which two employees work for a manager in a multitasking (dual-task) environment for two periods. In both experiments, we manipulate task complexity at two levels (a setting with two demanding tasks versus a setting with two less demanding tasks), because accounting task settings can vary strongly in complexity and complexity is one of the most important determinants of performance (Bonner & Sprinkle, 2002). After each period, the managers receive per employee a set of performance measures on each task and they reward the employees with a bonus. The two experiments differ in their management control environment. In the first experiment the managers are required to allocate one bonus pool for overall performance on the two tasks. In the second experiment managers can allocate a separate bonus pool for performance on each task separately.

Our findings indicate that if employees' allocation of effort across tasks deviates from the firm-preferred equal allocation of effort, this leads to lower employee overall performance on both tasks (H1). This is consistent with the findings of Hannan et al. (2013). We further demonstrate that the negative effect of distorted allocation of effort across tasks on overall employee performance on both tasks is less negative for complex tasks than for simple tasks (H2). This is the case because the extra time dedicated to a single task caused by the distorted allocation of effort allows employees to develop and test different strategies in order to solve the task. Finding and applying a better strategy is especially worthwhile for handling complex tasks (Locke & Latham, 1990). Our results further indicate that in the first period, employees will focus on the task for which they have the highest skills (H3a) and employees will focus on the easiest task in order to perform at least well on that task (H3b), consistent with self-affirmation theory (Steele, 1988). This behavior leads to a distorted allocation of effort across tasks. We further demonstrate that managers' bonus allocation is in line with the RPI they receive about the employees (H4a), but we do not find evidence of managers' that take into account employees deviations from the firm-preferred effort allocation across tasks in their bonus allocation (H4b). Nonetheless we find that the management control system consisting of

a discretionary bonus pool that accompanies the provision of detailed RPI is able to redirect employees' effort towards firm-preferred effort allocations across tasks. Consistent with social comparison theory and expectancy theory (Festinger, 1954; Vroom, 1964), we find that after the provision of RPI and the performance evaluation, employees reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague (H5).

We thereby hope to contribute to academic research related to the effect of RPI and distortions in effort allocations across tasks on employee performance in a multi-tasking context (Brüggen, 2011; Brüggen & Moers, 2007; Eichhorn, 2016; Farkas, 2013; Hannan et al., 2013). More specifically we demonstrate how discretionary bonus pools accompanying detailed RPI may affect employees' effort allocations across tasks (H5) and ultimately employee performance (H1 & 2). Our findings seem particularly useful for practice as well. Firms commonly provide RPI to their employees, but earlier research suggests that the provision of this RPI leads to distortions in effort allocation and lower employee performance (Farkas, 2013; Hannan et al., 2013). Our results indicate that RPI can lead to less distortion in effort allocation and higher employee performance. This is presumably caused by the fact that in our experiments RPI is accompanied with a discretionary bonus pool that allows the manager to direct employees' focus towards organizational preferences and objectives. However, further research is needed, as our experiments only focus on settings in which a discretionary bonus pool is present. It seems interesting to investigate the difference in employee performance and employee distortion depending on whether the firm only provides RPI or whether the firm accompanies RPI with a discretionary bonus pool. Farkas (2013), for instance, investigated the joint effect of financial incentives and RPI. Her results indicate that financial incentives are not effective at directing employee effort across tasks or in improving performance. Farkas (2013) further demonstrates that RPI has a negative effect on employees' effort allocation and performance in the presence of financial incentives. However, this result may be affected by the fact that employees did not believe they were eligible for compensation (Farkas, 2013). A discretionary bonus pool that does not contain hard-to-reach performance targets may be a better tool in promoting organizational desirable performance. Especially in a multitasking environment, a discretionary system seems particularly useful, since a complete contract specifying all possible outcomes and related employee rewards would be impossible (Gibbs et al., 2004; MacLeod and Parent, 1999; Rajan & Reichelstein, 2006).

Another area for future research lies in investigating the effect of the number of employees included in the RPI performance overviews. Most papers focus on a setting with 5 employees, whereas our paper focuses on a setting with 2 employees (Farkas, 2013; Hannan et al., 2013; Kramer et al., 2013). Social comparison concerns, competitive behavior amongst peers, employee motivation and employee effort decrease if the number of peers increases (Eichhorn, 2016; Garcia & Tor, 2009). There can only be one person that outperforms all the others and that flourishes by the RPI. All the others fall behind and can only distinguish themselves by focusing on one of the tasks of the multi-task setting. It seems probable that whether the predictions by self-affirmation theory (Steele, 1988) ('employees will distort their effort allocation to a single task in order to affirm their competence at least in one area') or social comparison theory (Festinger, 1954) ('employee will try to outperform their employees on all

areas') hold depends on the number of people featuring in the RPI performance overview. If the RPI only consists of two persons, the bottom performer can only substantially distinguish himself by improving his performance on both tasks. Focusing on one task will not increase his status, as he still remains the bottom performer. However, in a setting with 5 peers the top performer may be out of reach, but an average or bottom performer can still gain social status by outperforming his fellow average or bottom peers by focusing on a single task.

Another area for future research lies in the investigation of detailed RPI in a multitasking context. Most studies on RPI focus on rank orderings (Kramer et al., 2016 being an exception in a single-task setting). However, with rank orderings employees are unable to identify how close their performance is to the performance of others. Especially when employees can see that peer performance is close to one's own performance, social comparison leads to assimilation (the belief that one can achieve the same performance as the comparison target) (Eichhorn, 2016; Smith, 2000), which should lead to fierce competition and a willingness to perform well on all tasks (Festinger, 1954).

Lastly, the allocation of the discretionary bonus in our experiment was not publicly available to the employees. An employee did only know how many bonus points he received and the total amount of bonus points that could be divided. He would not know how many bonus points his colleague received. This is in line with pay secrecy practices in many firms in the US and abroad (Bamberger & Belogolovsky, 2010; Belogolovsky & Bamberger, 2014; Colella et al., 2007; Day, 2007; Thompson & Pronsky, 1975). However, previous literature indicates how the transparency of the bonus allocation affects how managers allocate a bonus (Bol et al., 2016). It seems likely that a public, transparent bonus attached to overall employee performance and compliance to firm-preferred effort allocations across tasks may serve as a tool for social distinction amongst employees. It then encourages employees to focus on overall performance on all tasks, which makes a strategy to focus on a single task obsolete, as this behavior will not (to a smaller extent) be rewarded by the manager. A similar effect was obtained in Eichhorn (2016) who demonstrates that unidimensional RPI indicating overall employee performance positively affects employee performance, whereas multidimensional RPI indicating employee performance on each separate task does not affect employee performance.

Our study is subject to some limitations as well, mainly related to the design of the second experiment. In our second experiment managers did not receive a performance overview containing employee performance on both tasks. They first received a performance overview containing RPI for task 1 and they then allocated a bonus for task 1 before they saw a performance overview containing RPI for task 2 and allocated a bonus for task 2. In such a context it is hard for a manager to judge the extent to which employees distort their effort allocations across tasks. Therefore, it is not that surprising that we do not find support for our hypothesis 4b (ceteris paribus, managers' bonus allocation will be lower when employees performance deviates from the firm-preferred effort allocation) in experiment 2. If the managers would have received a performance overview containing RPI for both tasks, it seems more likely that managers would take employees' distortions in effort allocation into account in their bonus allocation. However, in our first experiment managers had access to RPI for both tasks and we do not find support for hypothesis 4b in that experiment either. Finally, the order in which managers evaluated performance on each separate task in experiment 2 was fixed.

Managers always first evaluated performance on task 1 and then performance on task 2 and employees were aware of this order. In such a context employees may be inclined to overly focus on the first task in order to create a halo effect in their performance evaluation (Balzer & Sulsky, 1992).

General Conclusion

In this general conclusion, we first summarize the dissertation's main contributions to the literature, we then provide its implications for practice and we end with some of the limitations and opportunities for future research.

Contribution to the literature

This dissertation contributes to the literature on managers' use of their decision-making power to evaluate and reward employees' performance subjectively (Ahn et al., 2010; Bailey et al., 2011; Bol, 2008; Maas et al., 2012). Chapter1 provides an overview of academic research on subjective performance evaluation. Chapters 2 to 4 examine how managers use their decision-making power depending on MCS design elements such as: manager-employee compensation inequality and managerial discretion to freely decide on the bonus size (chapter 2); the accuracy of the performance information on which managers base their evaluations and whether (or not) managers get the opportunity to write a justification on how they allocated employee bonuses (chapter 3), and relative performance information that measures performance of an employee in relation to that of one or more other employees (chapter 4). We use various experiments to investigate these issues. Previous academic literature indicates the importance of empirical studies on how (elements of) the MCS and information produced by the MCS affect the behavior and decisions of individuals (Sprinkle, 2003; Sprinkle & Williamson, 2007). This dissertation intends to contribute to this literature.

Our first experimental study examines how managerial discretion in determining employee bonus size and manager-employee compensation inequality influence the proportion of the bonus pool that a manager allocates to her employee (knowing that the manager can keep the rest for herself). This study contributes to the literature, documenting benefits and drawbacks of control systems with explicit incentives to mitigate employer opportunism (Fisher et al., 2005). Previous academic research suggests that a reduction in managerial discretion in employee bonus *size* is beneficial, as it causes a manager to act less opportunistically, leading to higher employee effort and advantageous employee bonuses (Fisher et al., 2005). We demonstrate that high managerial discretion over employee bonus *size* can be beneficial to employees in situations with high compensation inequality because, in the presence of compensation inequality, managers with high discretion take fairness to a larger extent into account in their bonus decision than manager with low managerial discretion. This study further contributes to the literature on subjective performance evaluation by indicating how social preferences (reciprocity and fairness concerns) influence managers' use of discretion in bonus allocations (Abernethy et al., 2013; Maas et al., 2012). More specifically, managers positively reciprocate employees' effort with higher bonus allocations and managers that are concerned with fairness provide higher bonuses to their employees. Finally, the study contributes to the academic literature on the effect of compensation inequality on trust or employee effort (Anderson et al., 2006; Brülhart & Usunier, 2012; Greiner et al., 2012; Hargreaves Heap et al., 2013; Johnson & Mislin, 2011; Smith, 2011), as this study finds no effect of compensation inequality on employee effort and trust.

Our second study researches how performance information accuracy and process accountability (whether managers get the opportunity to write a justification for their bonus allocation) influence managers' differentiation in bonus allocations between employees. This study adds to the performance evaluation literature by indicating that different elements of the management control system jointly influence evaluation behavior (Bol et al., 2016; Libby et al., 2004). This study further contributes to the literature on compressed rating behavior by providing more details about the underlying process leading to this rating behavior (Ahn et al., 2010; Bailey et al., 2011; Bol, 2011; Bol et al., 2016; Golman & Bhatia, 2011; Levin, 2003; MacLeod, 2003; Moers, 2005; Prendergast & Topel, 1993). We show that a combination of highly accurate performance information and process accountability increases managers' estimates of employees' procedural fairness perceptions and employee acceptance of differentiation in bonus allocations, which consequently leads to less compressed bonus allocations. Furthermore, we provide additional empirical evidence for the informativeness principle (Holmström, 1979) by demonstrating that accurate performance information increases differentiation in bonus allocations. Finally, the study adds to the literature on accountability in performance evaluations by showing that the effect of accountability differs depending on the level of accuracy of the performance measurement system (Kennedy, 1993; Mero et al., 2003; Mero et al., 2007; Shore & Tashchian, 2002).

Our third study investigates how task complexity and distortion in effort allocation across tasks away from the firm-preferred effort allocation affect employee performance in a multitasking context. Additionally, the study examines how a management control system providing detailed relative performance information (RPI) and a discretionary bonus affects employees' effort allocation across tasks. This study contributes to the literature investigating effort allocation concerns in a multitasking setting (Bonner & Sprinkle, 2002; Brüggen & Moers, 2007; Hannan et al., 2013, 2017; Hecht et al., 2012) by documenting how performance-contingent monetary incentives and detailed RPI can affect employees' allocations and levels of effort among various tasks (Bonner & Sprinkle, 2002; Hannan et al., 2013, 2017). Many studies focus on rank information instead of detailed RPI (Farkas, 2013; Hannan et al., 2013), or do not provide performance-dependent pay (Hannan et al., 2013; Kramer et al., 2016). Our study demonstrates that, in the first period, employees will focus on the task for which they have the highest skills or find easiest in order to perform at least well on that task. Furthermore, we show that managers' bonus allocation is in line with the detailed RPI and that after the provision of RPI and the performance evaluation, employees reallocate their time across tasks such that they focus more (less) on tasks for which they under(out)performed relative to their colleague, consistent with social comparison theory and expectancy theory (Festinger, 1954; Vroom, 1964). Our findings further indicate that a distorted allocation of effort across tasks leads to lower employee overall performance, consistent with Hannan et al. (2013). However, the effect of distorted allocation of effort across tasks on overall employee performance on both tasks is less negative in a work environment consisting of complex tasks than in a work environment with simple tasks.

Implications for practice

Given the omnipresence of subjectivity in bonus plans, this dissertation on the use and consequences of subjective performance evaluation in supervisor-employee relationships is highly relevant for practice (WorldatWork & Vivient Consulting, 2012).

Our first study indicates that firms should consider the level of manager-employee compensation inequality when deciding on the level of managerial discretion. Depending on the extent of managerial discretion in combination with compensation inequality, managers' mindset towards fairness changes such that they balance concerns for long-term self-interest versus fairness in a different way. More specifically, in the presence of compensation inequality, managers with high discretion take fairness to a larger extent into account in their bonus decision than managers with low managerial discretion. This finding is important, given the attention manager-employee compensation inequality receives in the current debate on performance rewards (Anderson et al., 2006; Bloomberg, 2014; Brülhart & Usunier, 2012; Greiner et al., 2012; Guo et al., 2017; Hargreaves Heap et al., 2013; Johnson & Mislin, 2011; Reuters, 2013, 2014; Smith, 2011; The Economist, 2009).

Our second study pinpoints that firms should consider process accountability, next to information accuracy, when allocating bonuses. Highly accurate performance information leads directly to more differentiation in bonus allocations, which is desirable for organizations. After all, more differentiation leads to advantageous outcomes for organizations such as higher employee incentives, productivity, performance and performance improvement and better recognition of which employees might be eligible for promotion or additional training (Ahn et al., 2010; Baker et al., 1988; Bol, 2011; Golman & Bhatia, 2012). However, extra investments in more accurate performance information are often expensive. Our results indicate that adding a system of process accountability will lead to even more differentiation than the impact of accuracy of information in itself. As such, if organization make the costly investments related to higher data quality, they might as well provide managers with process accountability for a better - more differentiated – bonus allocation: process accountability will encourage managers to spend time and effort on information processing and will strengthen the pay-for-performance relationship. However, if organization do not invest in more accurate performance information, process accountability is not worthwhile, as it wastes the manager's valuable productive time without altering the degree of differentiation in the bonus allocation.

Our third study indicates that relative performance information (RPI) and bonus allocations attached to detailed RPI are able to steer employees' effort allocations across tasks towards firm preferences, leading to higher employee performance. Time management is an important part of employees' job (Adler & Benbunan-Fich, 2012; Farkas, 2013; Gonzalez & Mark, 2004; Kushleyeva et al., 2005; Payne et al., 2007). If organization can design management control systems that motivate the appropriate level and allocation of effort across tasks according to what the firm values based on its strategy, objectives or just the direct revenues related to each task, this will obviously impact firm's performance positively (Farkas, 2013).

Limitations and opportunities for future research

Although this dissertation provides contributions to the literature and practice, as all academic research, it is subject to a number of limitations. However, at the same time, these

limitations provide interesting avenues for future research. In all experiments, we research how solely one or two MCS design elements, in isolation, (jointly) affect how managers use their decision-making power in a particular, simplified performance evaluation setting. In practice many more design elements may impact managers' use of their managerial discretion. It can be questioned whether stylized, computer-based settings mimic all relevant aspects of performance evaluations in reality.

For example, in our second study, managers received the possibility to justify their bonus allocation, but solely to their employees. It seems likely that managers' rating behavior will be affected by whether they justify solely towards their employees, towards their superior or towards both the superior and the employees. In our first study, in contrast, managers did not receive the possibility to justify their bonus allocation. However, a manager's justification towards the employee could possibly have reduced conflicts, misinterpretations and faulty attributions, which could have motivated employee effort. An employee's justification of his behavior towards the manager or a manager's justification of his behavior towards its superior could have reduced managers' opportunistic behavior as well.

Next, the management control system in our experiments was imposed on the participants, was already installed and was not subject to discussion or alteration. However in reality the MCS might be questioned from time to time and both managers and employees may be involved in the design of the MCS (empowerment). Alternative MCS may lead to different evaluating behavior of managers. Furthermore, in reality managers and employees can choose to join a company with a MCS that fits their preferences or they can choose to leave a company that does not (longer) fit their preferences. This choice may affect their fairness perceptions or behavior. In the experiments, managers were not able to fire employees and employees were not able to resign. In practice, employees' potential threat to leave the company in case of insufficient or unfair bonus allocations and the managers' potential threat of dismissal in case of insufficient employee effort or in case of distortion in employee effort away from firm preferences, put a cap on employees' and managers' opportunistic behavior in practice.

Finally, all three experiments were computer-based, anonymous experiments, which rules out face-to-face conflicts and lasting, negative consequences. Furthermore, in the computer-based experiments, most performance information, payoffs and decisions were rather explicit and transparent. In reality this process is more ambiguous and secret, affecting managers' performance evaluations and employees behavior as well.

This interesting topic provides many more opportunities for future research. I am looking forward to tackle these issues in future research projects!

Bibliography

Abernethy, M. A., Hung, C. Y., & van Lent, L. (2013). Status and Discretionary Bonus Payments: Evidence from a Chinese Hospital. Working paper.

Adams, J. S. (1963). Towards an understanding of inequity. *The Journal of Abnormal and Social Psychology*, 67(5), 422.

Adams, J. S. (1965). Inequity in social exchange. *Advances in Experimental Social Psychology*, 2, 267-299.

Adler, R. F., & Benbunan-Fich, R. (2012). Juggling on a high wire: Multitasking effects on performance. *International Journal of Human-Computer Studies*, 70(2), 156-168.

Ahn, T. S., Hwang, I., & Kim, M. I. (2010). The impact of performance measure discriminability on ratee incentives. *The Accounting Review*, 85(2), 389-417.

Anderson, L. R., Mellor, J. M., & Milyo, J. (2006). Induced heterogeneity in trust experiments. *Experimental Economics*, 9(3), 223-235.

Ashton, R. H. (1990). Pressure and performance in accounting decision settings: Paradoxical effects of incentives, feedback, and justification. *Journal of Accounting Research*, 28, 148-180.

Axelrod, R. (1980). Effective choice in the prisoner's dilemma. *Journal of Conflict Resolution*, 24(1), 3-25.

Axelrod, R. (1981). The emergence of cooperation among egoists. *American Political Science Review*, 75(2), 306-318.

Axelrod, R. (1986). An evolutionary approach to norms. *American political science review*, 80(4), 1095-1111.

Bailey, B. P., & Konstan, J. A. (2006). On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state. *Computers in Human Behavior*, 22(4), 685-708.

Bailey, W. J., Hecht, G., & Towry, K. L. (2011). Dividing the pie: The influence of managerial discretion extent on bonus pool allocation. *Contemporary Accounting Research*, 28(5), 1562-1584.

Baiman, S., & Rajan, M. V. (1995). The informational advantages of discretionary bonus schemes. *The Accounting Review*, 70(4), 557-579.

Baker, G. P., Jensen, M. C., & Murphy, K. J. (1988). Compensation and incentives: Practice vs. theory. *The Journal of Finance*, 43(3), 593-616.

Baker, G., Gibbons, R., & Murphy, K. J. (1994). Subjective performance measures in optimal incentive contracts. *The Quarterly Journal of Economics*, 109(4), 1125-1156.

Balzer, W. K., & Sulsky, L. M. (1992). Halo and performance appraisal research: A critical examination. *Journal of Applied Psychology*, 77(6), 975.

Bamberger, P., & Belogolovsky, E. (2010). The impact of pay secrecy on individual task performance. *Personnel Psychology*, 63(4), 965-996.

Banker, R. D., & Datar, S. M. (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research*, 27(1), 21-39.

Banker, R. D., Chang, H., & Pizzini, M. J. (2004). The balanced scorecard: Judgmental effects of performance measures linked to strategy. *The Accounting Review*, 79(1), 1-23.

BBC. (2013). Swiss vote no to capping bosses' pay at 12 times lowest paid. http://www.bbc.com/news/business-25076879

Belogolovsky, E., & Bamberger, P. A. (2014). Signaling in secret: Pay for performance and the incentive and sorting effects of pay secrecy. *Academy of Management Journal*, 57(6), 1706-1733.

Berg, J., Dickhaut, J., & McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1), 122-142.

Bernardin, H. J., & Villanova, P. (1986). Performance appraisal. In *Generalizing from Laboratory to Field Settings*, edited by E. A. Locke. Lexington, MA: Lexington Books D.C. Heath and Company.

Biernat, M., & Sesko, A. K. (2013). Evaluating the contributions of members of mixed-sex work teams: Race and gender matter. *Journal of Experimental Social Psychology*, 49(3), 471-476.

Bies, R. J., & Shapiro, D. L. (1988). Voice and justification: Their influence on procedural fairness judgments. *Academy of Management Journal*, 31(3), 676-685.

Bloomberg. (2013). CEO-to-Worker Pay-Ratio Disclosure Proposed by Divided SEC. http://www.bloomberg.com/news/2013-09-17/ceo-to-worker-pay-ratio-disclosure-proposal-to-be-issued-by-sec.html

Bloomberg. (2014). Staples Shareholders Vote Against Executive-Compensation Plan. http://www.bloomberg.com/news/articles/2014-06-03/staples-shareholders-vote-against-executive-compensation-plan

Bol, J. C. (2008). Subjectivity in Compensation Contracting. *Journal of Accounting Literature*, 27, 1-24.

Bol, J. C. (2011). The determinants and performance effects of managers' performance evaluation biases. *The Accounting Review*, 86(5), 1549-1575.

Bol, J. C., & Smith, S. D. (2011). Spillover effects in subjective performance evaluation: Bias and the asymmetric influence of controllability. *The Accounting Review*, 86(4), 1213-1230.

Bol, J. C., Kramer, S., & Maas, V. S. (2016). How control system design affects performance evaluation compression: The role of information accuracy and outcome transparency. *Accounting, Organizations and Society*, 51, 64-73.

Bolton, G. E., & Ockenfels, A. (2000). ERC: A theory of equity, reciprocity, and competition. *The American Economic Review*, 90 (1), 166-193.

Bommer, W. H., Johnson, J. L., Rich, G. A., Podsakoff, P. M., & MacKenzie, S. B. (1995). On the interchangeability of objective and subjective measures of employee performance: A meta-analysis. *Personnel Psychology*, 48(3), 587-605.

Bonner, S. E. (1999). Judgment and decision-making research in accounting. *Accounting Horizons*, 13(4), 385-398.

Bonner, S. E. (2008). Judgment and decision making in accounting. Prentice Hall.

Bonner, S. E., & Sprinkle, G. B. (2002). The effects of monetary incentives on effort and task performance: theories, evidence, and a framework for research. *Accounting, Organizations and Society*, 27(4), 303-345.

Bonner, S. E., Hastie, R., Sprinkle, G. B., & Young, S. M. (2000). A review of the effects of financial incentives on performance in laboratory tasks: Implications for management accounting. *Journal of Management Accounting Research*, 12(1), 19-64.

Borst, J. P., Taatgen, N. A., & van Rijn, H. (2010). The problem state: A cognitive bottleneck in multitasking. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 36(2), 363.

Brüggen, A. (2011). Ability, career concerns, and financial incentives in a multi-task setting. *Journal of Management Accounting Research*, 23(1), 211-229.

Brüggen, A., & Moers, F. (2007). The role of financial incentives and social incentives in multi-task settings. *Journal of Management Accounting Research*, 19(1), 25-50.

Brülhart, M., & Usunier, J. C. (2012). Does the trust game measure trust?. *Economics Letters*, 115(1), 20-23.

Brutus, S. (2010). Words versus numbers: A theoretical exploration of giving and receiving narrative comments in performance appraisal. *Human Resource Management Review*, 20(2), 144-157.

Buckless, F. A., & Ravenscroft, S. P. (1990). Contrast coding: A refinement of ANOVA in behavioral analysis. *The Accounting Review*, 65(4), 933–945.

Budde, J. (2007). Performance measure congruity and the balanced scorecard. *Journal of Accounting Research*, 45(3), 515-539.

Bull, C. (1987). The existence of self-enforcing implicit contracts. *The Quarterly Journal of Economics*, 102(1), 147-159.

Burney, L. L., Henle, C. A., & Widener, S. K. (2009). A path model examining the relations among strategic performance measurement system characteristics, organizational justice, and extra-and in-role performance. *Accounting, Organizations and Society*, 34(3), 305-321.

Buser, T., & Peter, N. (2012). Multitasking. *Experimental Economics*, 15(4), 641-655.

Campbell, D. J. (1988). Task complexity: A review and analysis. *Academy of Management Review*, 13(1), 40-52.

Cardinaels, E., & van Veen-Dirks, P. M. (2010). Financial versus non-financial information: The impact of information organization and presentation in a Balanced Scorecard. *Accounting, Organizations and Society*, 35(6), 565-578.

Castilla, E. J. (2008). Gender, race, and meritocracy in organizational careers. *American Journal of Sociology*, 113(6), 1479-1526.

Castilla, E. J. (2015). Accounting for the gap: A firm study manipulating organizational accountability and transparency in pay decisions. *Organization Science*, 26(2), 311-333.

Chenhall, R. H. (2003). Management control systems design within its organizational context: findings from contingency-based research and directions for the future. *Accounting, Organizations and Society*, 28(2), 127-168.

Choi, J., Hecht, G. W., & Tayler, W. B. (2012). Lost in translation: The effects of incentive compensation on strategy surrogation. *The Accounting Review*, 87(4), 1135-1163.

Choi, J., Hecht, G., Tafkov, I. D., & Towry, K. L. (2015). Vicarious learning under implicit contracts. *The Accounting Review*, 91(4), 1087-1108.

Christ, M. H., Emett, S. A., Tayler, W. B., & Wood, D. A. (2016). Compensation or feedback: Motivating performance in multidimensional tasks. *Accounting, Organizations and Society*, 50, 27-40.

Christ, M. H., Sedatole, K. L., & Towry, K. L. (2012). Sticks and carrots: The effect of contract frame on effort in incomplete contracts. *The Accounting Review*, 87(6), 1913-1938.

Ciriolo, E. (2007). Inequity aversion and trustees' reciprocity in the trust game. *European Journal of Political Economy*, 23(4), 1007-1024.

Colella, A., Paetzold, R. L., Zardkoohi, A., & Wesson, M. J. (2007). Exposing pay secrecy. *Academy of Management Review*, 32(1), 55-71.

Coletti, A. L., Sedatole, K. L., & Towry, K. L. (2005). The effect of control systems on trust and cooperation in collaborative environments. *The Accounting Review*, 80(2), 477-500.

Colquitt, J. A. (2001). On the dimensionality of organizational justice: a construct validation of a measure. *Journal of Applied Psychology*, 86(3), 386.

Cox, J. C. (2004). How to identify trust and reciprocity. *Games and Economic Behavior*, 46(2), 260-281.

Cronqvist, H., & Fahlenbrach, R. (2013). CEO contract design: How do strong principals do it?. *Journal of Financial Economics*, 108(3), 659-674.

Daily Mail. (2013). Swiss vote to impose world's strictest rules on executive pay after public outcry over fat cat bonuses. http://www.dailymail.co.uk/news/article-2287432/Swiss-vote-impose-worlds-strictest-rules-executive-pay-public-outcry-fat-cat-bonuses.html

Day, N. E. (2007). An investigation into pay communication: is ignorance bliss?. *Personnel Review*, 36(5), 739-762.

De Standaard. (2013a). Bpost en Labille op ramkoers over loon CEO. http://www.standaard.be/cnt/dmf20131217_00891478

De Standaard. (2013b). Loon voor opvolger Bellens blijft voorwerp van discussie. http://www.standaard.be/cnt/dmf20131210_00880595

Dickinson, T. L. (1993). Attitudes about performance appraisal. *Personnel Selection and Assessment: Individual and Organizational Perspectives*, 141-161.

Ding, S., & Beaulieu, P. (2011). The Role of Financial Incentives in Balanced Scorecard-Based Performance Evaluations: Correcting Mood Congruency Biases. *Journal of Accounting Research*, 49(5), 1223-1247.

Dossett, D. L., & Greenberg, C. I. (1981). Goal setting and performance evaluation: An attributional analysis. *Academy of Management Journal*, 24(4), 767-779.

Du, F., Tang, G., & Young, S. M. (2012). Influence activities and favoritism in subjective performance evaluation: evidence from Chinese state-owned enterprises. *The Accounting Review*, 87(5), 1555-1588.

Duarte, N. T., Goodson, J. R., & Klich, N. R. (1994). Effects of dyadic quality and duration on performance appraisal. *Academy of Management Journal*, 37(3), 499-521.

Duggan, G. B., Johnson, H., & Sørli, P. (2013). Interleaving tasks to improve performance: Users maximise the marginal rate of return. *International Journal of Human-Computer Studies*, 71(5), 533-550.

Dulebohn, J. H., & Ferris, G. R. (1999). The role of influence tactics in perceptions of performance evaluations' fairness. *Academy of Management Journal*, 42(3), 288-303.

Dur, R., & Glazer, A. (2007). Optimal contracts when a worker envies his boss. *The Journal of Law, Economics, & Organization*, 24(1), 120-137.

Dux, P. E., Tombu, M. N., Harrison, S., Rogers, B. P., Tong, F., & Marois, R. (2009). Training improves multitasking performance by increasing the speed of information processing in human prefrontal cortex. *Neuron*, 63(1), 127-138.

Eichhorn, N. (2016). Relative Performance Information and Financial Incentives in Multidimensional Task Settings–A Conceptual and Experimental Analysis of Effects on Performance and Attention towards Task Dimensions. *Junior Management Science*, 1(1), 100-137.

Elvira, M., & Town, R. (2001). The effects of race and worker productivity on performance evaluations. *Industrial Relations: A Journal of Economy and Society*, 40(4), 571-590.

Ewing, R. (2016). An Experimental Investigation of the Effects of Contract Frame and Discretion in Performance Evaluation on Effort. Theses and Dissertations—Accountancy. University of Kentucky Uknowledge.

Falk, A., & Fischbacher, U. (2006). A theory of reciprocity. *Games and Economic Behavior*, 54(2), 293-315.

Farkas, M. J. (2013). *Multi-Task Setting Involving Simple and Complex Tasks: An Exploratory Study of Employee Motivation*. University of South Florida.

Fehr, E., & Gächter, S. (2000). Fairness and retaliation: The economics of reciprocity. *The journal of Economic Perspectives*, 14(3), 159-181.

Fehr, E., & Schmidt, K. M. (1999). A theory of fairness, competition, and cooperation. *The Quarterly Journal of Economics*, 114(3), 817-868.

Fehr, E., & Schmidt, K. M. (2004). Fairness and Incentives in a Multi-task Principal–Agent Model. *The Scandinavian Journal of Economics*, 106(3), 453-474.

Fehr, E., Gächter, S., & Kirchsteiger, G. (1997). Reciprocity as a contract enforcement device: Experimental evidence. *Econometrica: Journal of the Econometric Society*, 833-860.

Ferreira, A., & Otley, D. (2009). The design and use of performance management systems: An extended framework for analysis. *Management Accounting Research*, 20(4), 263-282.

Ferris, G. R., Munyon, T. P., Basik, K., & Buckley, M. R. (2008). The performance evaluation context: Social, emotional, cognitive, political, and relationship components. *Human Resource Management Review*, 18(3), 146-163.

Festinger, L. (1954). A theory of social comparison processes. *Human Relations*, 7(2), 117-140.

Fischbacher, U. (2007). z-Tree: Zurich toolbox for ready-made economic experiments. *Experimental Economics*, 10(2), 171-178.

Fisher, J. G., Maines, L. A., Peffer, S. A., & Sprinkle, G. B. (2005). An experimental investigation of employer discretion in employee performance evaluation and compensation. *The Accounting Review*, 80(2), 563-583.

Franco-Santos, M., Lucianetti, L., & Bourne, M. (2012). Contemporary performance measurement systems: A review of their consequences and a framework for research. *Management Accounting Research*, 23(2), 79-119.

Frey, B. S. (2007). Awards as compensation. *European Management Review*, 4(1), 6-14.

Frink, D. D., & Ferris, G. R. (1999). The moderating effects of accountability on the conscientiousness-performance relationship. *Journal of Business and Psychology*, 13(4), 515-524.

Garcia, S. M., & Tor, A. (2009). The N-effect: More competitors, less competition. *Psychological Science*, 20(7), 871-877.

Ghosh, D., & Lusch, R. F. (2000). Outcome effect, controllability and performance evaluation of managers: some field evidence from multi-outlet businesses. *Accounting, Organizations and Society*, 25(4), 411-425.

Gibbins, M., & Newton, J. D. (1994). An empirical exploration of complex accountability in public accounting. *Journal of Accounting Research*, 32 (4),165-186.

Gibbs, M., Merchant, K. A., Stede, W. A. V. D., & Vargus, M. E. (2004). Determinants and effects of subjectivity in incentives. *The Accounting Review*, 79(2), 409-436.

Gillie, T., & Broadbent, D. (1989). What makes interruptions disruptive? A study of length, similarity, and complexity. *Psychological Research*, 50(4), 243-250.

Giraud, F., Langevin, P., & Mendoza, C. (2008). Justice as a rationale for the controllability principle: A study of managers' opinions. *Management Accounting Research*, 19(1), 32-44.

Golman, R., & Bhatia, S. (2012). Performance evaluation inflation and compression. *Accounting, Organizations and Society*, 37(8), 534-543.

González, V. M., & Mark, G. (2004). Constant, constant, multi-tasking craziness: managing multiple working spheres. In *Proceedings of the SIGCHI conference on Human factors in computing systems* (pp. 113-120). ACM.

Govindarajan, V. (1984). Appropriateness of accounting data in performance evaluation: an empirical examination of environmental uncertainty as an intervening variable. *Accounting, Organizations and Society*, 9(2), 125-135.

Greiner, B., Ockenfels, A., & Werner, P. (2012). The dynamic interplay of inequality and trust—an experimental study. *Journal of Economic Behavior & Organization*, 81(2), 355-365.

Guo, L., Libby, T., & Liu, X. K. (2017). The effects of vertical pay dispersion: Experimental evidence in a budget setting. *Contemporary Accounting Research*, 34(1), 555-576.

Guymon, R. N., (2008). The Effect of Explanations and Monetary Incentives on Effort Allocation Decisions . AAA 2009 Management Accounting Section (MAS) Meeting Paper. Available at SSRN: https://ssrn.com/abstract=1193142 or http://dx.doi.org/10.2139/ssrn.1193142

Hales, J., & Williamson, M. G. (2010). Implicit employment contracts: The limits of management reputation for promoting firm productivity. *Journal of Accounting Research*, 48(1), 147-176.

Hannan, R. L., McPhee, G. P., Newman, A. H. & Tafkov, I., (2017). Designing a Performance Feedback System in a Multi-Task Environment: Relative Performance Information Detail Level and Temporal Aggregation in a Multi-Task Environment (February 3, 2017). Available at SSRN: https://ssrn.com/abstract=2911072

Hannan, R. L., McPhee, G. P., Newman, A. H., & Tafkov, I. D. (2013). The effect of relative performance information on performance and effort allocation in a multi-task environment. *The Accounting Review*, 88(2), 553-575.

Hargreaves Heap, S. P. H., Tan, J. H., & Zizzo, D. J. (2013). Trust, inequality and the market. *Theory and Decision*, 74(3), 311-333.

Hartmann, F., & Slapničar, S. (2009). How formal performance evaluation affects trust between superior and subordinate managers. *Accounting, Organizations and Society*, 34(6), 722-737.

Hartmann, F., & Slapničar, S. (2012). The perceived fairness of performance evaluation: The role of uncertainty. *Management Accounting Research*, 23(1), 17-33.

Hartmann, F., Naranjo-Gil, D., & Perego, P. (2010). The effects of leadership styles and use of performance measures on managerial work-related attitudes. *European Accounting Review*, 19(2), 275-310.

Harvard Business Review. (2009). Why Sky-High CEO Pay Is Bad Business. https://hbr.org/2009/06/why-high-ceo-pay-is-bad-business/

Hecht, G., Tafkov, I., & Towry, K. L. (2012). Performance spillover in a multitask environment. *Contemporary Accounting Research*, 29(2), 563-589.

Heneman, R. L., & Wexley, K. N. (1983). The effects of time delay in rating and amount of information observed on performance rating accuracy. *Academy of Management Journal*, 26(4), 677-686.

Hermans, T., Cools, M., & Van den Abbeele, A. (2013). Subjective performance measurement: a literature review. *Rev Bus Econ Lit*, 58(4), 308-342.

Hermans, T., Cools, M., & Van den Abbeele. (2017). A. Subjective performance evaluation: the role of information accuracy and accountability –working paper.

Hoffman, V. B., & Patton, J. M. (1997). Accountability, the dilution effect, and conservatism in auditors' fraud judgments. *Journal of Accounting Research*, 35(2), 227-237.

Hogan, E. A. (1987). Effects of prior expectations on performance ratings: A longitudinal study. *Academy of Management Journal*, 30(2), 354-368.

Holmström, B. (1979). Moral hazard and observability. *The Bell Journal of Economics*, 74-91.

Holmström, B., & Milgrom, P. (1991). Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7, 24-52.

Höppe, F., & Moers, F. (2011). The choice of different types of subjectivity in CEO annual bonus contracts. *The Accounting Review*, 86(6), 2023-2046.

Horngren, C. T., Datar, S. M., & Rajan, M. V. (2015). Cost Accounting: A Managerial Emphasis. Fifteenth Edition. Pearson education.

Humphreys, K. A., & Trotman, K. T. (2011). The balanced scorecard: The effect of strategy information on performance evaluation judgments. *Journal of Management Accounting Research*, 23(1), 81-98.

Indjejikian, R. J., & Matějka, M. (2011). Accounting decentralization and performance evaluation of business unit managers. *The Accounting Review*, 87(1), 261-290.

Itoh, H. (2004). Moral hazard and other-regarding preferences. *The Japanese Economic Review*, 55(1), 18-45.

Ittner, C. D., Larcker, D. F., & Meyer, M. W. (2003). Subjectivity and the weighting of performance measures: Evidence from a balanced scorecard. *The Accounting Review*, 78(3), 725-758.

Ivancevich, J. M. (1983). Contrast effects in performance evaluation and reward practices. *Academy of Management Journal*, 26(3), 465-476.

Jacobs, R., & Kozlowski, S. W. (1985). A closer look at halo error in performance ratings. *Academy of Management Journal*, 28(1), 201-212.

Janssen, C. P., Brumby, D. P., Dowell, J., Chater, N., & Howes, A. (2011). Identifying optimum performance trade-offs using a cognitively bounded rational analysis model of discretionary task interleaving. *Topics in Cognitive Science*, 3(1), 123-139.

Jasmand, C., Blazevic, V., & de Ruyter, K. (2012). Generating sales while providing service: A study of customer service representatives' ambidextrous behavior. *Journal of Marketing*, 76(1), 20-37.

Johnson, N. D., & Mislin, A. A. (2011). Trust games: A meta-analysis. *Journal of Economic Psychology*, 32(5), 865-889.

Judge, T. A., & Ferris, G. R. (1993). Social context of performance evaluation decisions. *Academy of Management Journal*, 36(1), 80-105.

Kanagaretnam, K., Mestelman, S., Nainar, S. K., & Shehata, M. (2012). The impact of empowering investors on trust and trustworthiness. *Journal of Economic Psychology*, 33(3), 566-577.

Kane, J. S., Bernardin, H. J., Villanova, P., & Peyrefitte, J. (1995). Stability of rater leniency: Three studies. *Academy of Management Journal*, 38(4), 1036-1051.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard--measures that drive performance. *Harvard Business Review*, 70(1), 71-79.

Ke, B., Petroni, K., & Safieddine, A. (1999). Ownership concentration and sensitivity of executive pay to accounting performance measures: Evidence from publicly and privately-held insurance companies. *Journal of Accounting and Economics*, 28(2), 185-209.

Keeley, M. (1977). Subjective performance evaluation and person-role conflict under conditions of uncertainty. *Academy of Management Journal*, 20(2), 301-314.

Kennedy, J. (1993). Debiasing audit judgment with accountability: A framework and experimental results. *Journal of Accounting Research*, 31(2), 231-245.

Klimoski, R., & Inks, L. (1990). Accountability forces in performance appraisal. *Organizational Behavior and Human Decision Processes*, 45(2), 194-208.

Konig, C. J., Buhner, M., & Murling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance*, 18(3), 243-266.

Kramer, S., & Maas, V. S. (2014). Information Search and Information Processing in Subjective Performance Evaluation: Evidence from an Eye-Tracking Experiment. Available at SSRN: https://ssrn.com/abstract=2457941

Kramer, S., Maas, V. S., & Van Rinsum, M. (2016). Relative performance information, rank ordering and employee performance: A research note. *Management Accounting Research*, 33, 16-24.

Kreps, D. M., Milgrom, P., Roberts, J., & Wilson, R. (1982). Rational cooperation in the finitely repeated prisoners' dilemma. *Journal of Economic Theory*, 27(2), 245-252.

Krishnan, R., Luft, J. L., & Shields, M. D. (2005). Effects of accounting-method choices on subjective performance-measure weighting decisions: Experimental evidence on precision and error covariance. *The Accounting Review*, 80(4), 1163-1192.

Kushleyeva, Y., Salvucci, D. D., & Lee, F. J. (2005). Deciding when to switch tasks in time-critical multitasking. *Cognitive Systems Research*, 6(1), 41-49.

Lahno, B. (1995). Trust and strategic rationality. *Rationality and Society*, 7(4), 442-464.

Landy, F. J., Barnes, J. L., & Murphy, K. R. (1978). Correlates of Perceived Fairness and Accuracy of Performance Evaluation. *Journal of Applied Psychology*, 63(6), 751-754.

Lawler, E. E. (1967). The multitrait-multirater approach to measuring managerial job performance. *Journal of Applied Psychology*, 51(5p1), 369.

Lazarus, R. S. (1991). Emotion and adaptation. Oxford University Press on Demand.

Ledyard, J. 0. (1995). Public goods: A survey of experimental research. Handbook of Experimental Economics, Princeton University Press, Princeton: 111-194.

Lenton, P., & Mosley, P. (2011). Incentivising trust. *Journal of Economic Psychology*, 32(5), 890-897.

Lerner, J. S., & Tetlock, P. E. (1999). Accounting for the effects of accountability. *Psychological Bulletin*, 125(2), 255.

Leventhal, G. S. (1980). What should be done with equity theory?. In Social exchange (pp. 27-55). Springer US.

Levin, J. (2003). Relational incentive contracts. *American Economic Review*, 93(3), 835-857.

Libby, T., Salterio, S. E., & Webb, A. (2004). The balanced scorecard: The effects of assurance and process accountability on managerial judgment. *The Accounting Review*, 79(4), 1075-1094.

Lipe, M. G., & Salterio, S. (2002). A note on the judgmental effects of the balanced scorecard's information organization. *Accounting, Organizations and Society*, 27(6), 531-540.

Lipe, M. G., & Salterio, S. E. (2000). The balanced scorecard: Judgmental effects of common and unique performance measures. *The Accounting Review*, 75(3), 283-298.

Locke, E. A., & Latham, G. P. (1990). Work motivation and satisfaction: Light at the end of the tunnel. *Psychological Science*, 1(4), 240-246.

Lohman, C., Fortuin, L., & Wouters, M. (2004). Designing a performance measurement system: A case study. *European Journal of Operational Research*, 156(2), 267-286.

Luft, J. (2016). Cooperation and competition among employees: Experimental evidence on the role of management control systems. *Management Accounting Research*, 31, 75-85.

Luft, J. L. (1997). Fairness, ethics and the effect of management accounting on transaction costs. *Journal of Management Accounting Research*, 9, 199.

Luft, J., & Shields, M. D. (2003). Mapping management accounting: graphics and guidelines for theory-consistent empirical research. *Accounting, Organizations and Society*, 28(2), 169-249.

Maas, V. S., van Rinsum, M., & Towry, K. L. (2012). In Search of Informed Discretion: An Experimental Investigation of Fairness and Trust Reciprocity. *The Accounting Review*, 87(2), 617-644.

MacLeod, W. B. (2003). Optimal contracting with subjective evaluation. *American Economic Review*, 93(1), 216-240.

MacLeod, W. B., & Parent, D. (1999). Job characteristics, wages and the employment contract. *Review*, 81, 13-28.

Madjar, N., & Shalley, C. E. (2008). Multiple tasks' and multiple goals' effect on creativity: Forced incubation or just a distraction?. *Journal of Management*, 34(4), 786-805.

Malhotra, D., & Murnighan, J. K. (2002). The effects of contracts on interpersonal trust. *Administrative Science Quarterly*, 47(3), 534-559.

Masschelein, S., Cardinaels, E., & Van den Abbeele, A. (2012). ABC information, fairness perceptions, and interfirm negotiations. *The Accounting Review*, 87(3), 951-973.

Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.

McCabe, K. A., Rigdon, M. L., & Smith, V. L. (2003). Positive reciprocity and intentions in trust games. *Journal of Economic Behavior & Organization*, 52(2), 267-275.

McFarlin, D. B., & Sweeney, P. D. (1992). Distributive and procedural justice as predictors of satisfaction with personal and organizational outcomes. *Academy of Management Journal*, 35(3), 626-637.

Merchant, K. A., & Van der Stede, W. A. (2007). Management control systems: performance measurement, evaluation and incentives. Pearson Education.

Merchant, K. A., Chow, C. W., & Wu, A. (1995). Measurement, evaluation and reward of profit center managers: a cross-cultural field study. *Accounting, Organizations and Society*, 20(7-8), 619-638.

Mero, N. P., & Motowidlo, S. J. (1995). Effects of rater accountability on the accuracy and the favorability of performance ratings. *Journal of Applied Psychology*, 80(4), 517.

Mero, N. P., Guidice, R. M., & Brownlee, A. L. (2007). Accountability in a performance appraisal context: The effect of audience and form of accounting on rater response and behavior. *Journal of Management*, 33(2), 223-252.

Mero, N. P., Motowidlo, S. J., & Anna, A. L. (2003). Effects of accountability on rating behavior and rater accuracy. *Journal of Applied Social Psychology*, 33(12), 2493-2514.

Moers, F. (2005). Discretion and bias in performance evaluation: the impact of diversity and subjectivity. *Accounting, Organizations and Society*, 30(1), 67-80.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134-140.

Moore, D. A. (2007). Not so above average after all: When people believe they are worse than average and its implications for theories of bias in social comparison. *Organizational Behavior and Human Decision Processes*, 102(1), 42-58.

Morse, A., Nanda, V., & Seru, A. (2011). Are Incentive Contracts Rigged by Powerful CEOs?. *Journal of Finance*, 66(5), 1779-1821.

Murphy, K. J., & Oyer, P. (2003). Discretion in executive compensation contracts. Working paper, University of Southern California and Stanford University.

Murthy, U. S., & Schafer, B. A. (2011). The effects of relative performance information and framed information systems feedback on performance in a production task. *Journal of Information Systems*, 25(1), 159-184.

Ockenfels, A. (2015). Tit for Tat. *German Research*, 37(3), 12-15.

OECD. (2011a). An Overview of Growing Income Inequalities in OECD Countries: Main Findings. http://www.oecd.org/els/soc/49499779.pdf

OECD. (2011b). Board Practices: Incentives and Governing Risks. http://www.oecd.org/daf/ca/49081438.pdf

Otley, D. (1999). Performance management: a framework for management control systems research. *Management Accounting Research*, 10(4), 363-382.

Payne, S. J., Duggan, G. B., & Neth, H. (2007). Discretionary task interleaving: Heuristics for time allocation in cognitive foraging. *Journal of Experimental Psychology General*, 136(3), 370.

Pipino, L. L., Lee, Y. W., & Wang, R. Y. (2002). Data quality assessment. *Communications of the ACM*, 45(4), 211-218.

Prendergast, C. (1993). A Theory of" Yes Men". *The American Economic Review*, 83(4), 757-770.

Prendergast, C., & Topel, R. (1993). Discretion and bias in performance evaluation. *European Economic Review*, 37(2-3), 355-365.

Prendergast, C., & Topel, R. H. (1996). Favoritism in organizations. *Journal of Political Economy*, 104(5), 958-978.

Pulakos, E. D., & Wexley, K. N. (1983). The Relationship Among Perceptual Similarity, Sex, and Performance Ratings in Manager-subordinate Dyads. *Academy of Management Journal*, 26(1): 129-139.

Rabin, M. (1993). Incorporating Fairness into Game Theory and Economics. *The American Economic Review*, 83(5): 1281.

Rajan, M. V., & Reichelstein, S. (2006). Subjective performance indicators and discretionary bonus pools. *Journal of Accounting Research*, 44(3), 585-618.

Rajan, M. V., & Reichelstein, S. (2009). Objective versus subjective indicators of managerial performance. *The Accounting Review*, 84(1), 209-237.

Rand, D. G., Fudenberg, D., & Dreber, A. (2013). It's the thought that counts: The role of intentions in reciprocal altruism. Available at SSRN 2259407.

Reis, H. T., & Gruzen, J. (1976). On mediating equity, equality, and self-interest: The role of self-presentation in social exchange. *Journal of Experimental Social Psychology*, 12(5), 487-503.

Ren, D., Zhou, H., & Fu, X. (2009). A deeper look at gender difference in multitasking: Gender-specific mechanism of cognitive control. In *Natural Computation, 2009. ICNC'09. Fifth International Conference on* (Vol. 5, pp. 13-17). IEEE.

Reuters. (2013). COLUMN-Swiss outrage over executive pay sparks a movement in Europe. http://www.reuters.com/article/2013/11/15/gumbel-switzerland-idUSL2N0J01LI20131115

Reuters. (2014). Novartis keeps CEO pay steady, but incoming chairman pay drops. http://www.reuters.com/article/2014/01/29/us-novartis-results-compensation-idUSBREA0S0QS20140129

Roch, S. G. (2005). An investigation of motivational factors influencing performance ratings: Rating audience and incentive. *Journal of Managerial Psychology*, 20(8), 695-711.

Rubinstein, J. S., Meyer, D. E., & Evans, J. E. (2001). Executive control of cognitive processes in task switching. *Journal of Experimental Psychology: Human Perception and Performance*, 27(4), 763.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, 25(1), 54-67.

Salterio, S. E. (2015). Barriers to knowledge creation in management accounting research. *Journal of Management Accounting Research*, 27(1), 151-170.

Schmidt, K. M., & Schnitzer, M. (1995). The interaction of explicit and implicit contracts. *Economics letters*, *48*(2), 193-199.

Segal, E. (2004). Incubation in insight problem solving. *Creativity Research Journal*, 16(1), 141-148.

Shapiro, E. G. (1975). Effect of expectations of future interaction on reward allocations in dyads: Equity or equality. *Journal of Personality and Social Psychology*, 31(5), 873.

Shore, T. H., & Tashchian, A. (2002). Accountability forces in performance appraisal: Effects of self-appraisal information, normative information, and task performance. *Journal of Business and Psychology*, 17(2), 261-274.

Sloof, R., & Sonnemans, J. (2011). The interaction between explicit and relational incentives: An experiment. *Games and Economic Behavior*, 73(2), 573-594.

Smith, A. (2011). Income inequality in the trust game. *Economics Letters*, 111(1), 54-56.

Smith, R. H. (2000). Assimilative and contrastive emotional reactions to upward and downward social comparisons. *Handbook of social comparison: Theory and research*, 173-200.

Speier, C., Valacich, J. S., & Vessey, I. (1999). The influence of task interruption on individual decision making: An information overload perspective. *Decision Sciences*, 30(2), 337-360.

Speier, C., Vessey, I., & Valacich, J. S. (2003). The effects of interruptions, task complexity, and information presentation on computer-supported decision-making performance. *Decision Sciences*, 34(4), 771-797.

Sprinkle, G. B. (2003). Perspectives on experimental research in managerial accounting. *Accounting, Organizations and Society*, 28(2), 287-318.

Sprinkle G.B. & Williamson M.G. (2007). Experimental research in managerial accounting. *Handbook of Management Accounting Research*, Chapman, C.S., Hopwood, A.G. & Shields, M.D (Eds.) (pp. 415-444). Oxford, UK: Elsevier.

Steele, C. M. (1988). The psychology of self-affirmation: Sustaining the integrity of the self. *Advances in Experimental Social Psychology*, 21, 261-302.

Stoet, G., O'Connor, D. B., Conner, M., & Laws, K. R. (2013). Are women better than men at multi-tasking?. *BMC Psychology*, 1(1), 18.

Tafkov, I. D. (2013). Private and public relative performance information under different compensation contracts. *The Accounting Review*, 88(1), 327-350.

Tan, H. T., & Jamal, K. (2001). Do auditors objectively evaluate their subordinates' work?. *The Accounting Review*, 76(1), 99-110.

Tayler, W. B. (2010). The balanced scorecard as a strategy-evaluation tool: The effects of implementation involvement and a causal-chain focus. *The Accounting Review*, 85(3), 1095-1117.

Tenbrunsel, A. E., & Messick, D. M. (1999). Sanctioning systems, decision frames, and cooperation. *Administrative Science Quarterly*, 44(4), 684-707.

Tetlock, P. E., Skitka, L., & Boettger, R. (1989). Social and cognitive strategies for coping with accountability: conformity, complexity, and bolstering. *Journal of Personality and Social Psychology*, 57(4), 632.

The Economist. (2009). Attacking the corporate gravy train. http://www.economist.com/node/13726705

The Guardian. (2013). Switzerland votes against cap on executive pay. http://www.theguardian.com/world/2013/nov/24/switzerland-votes-against-cap-executive-pay

The US News. (2013). What We Can Learn From Switzerland's CEO Pay Cap Vote. http://www.usnews.com/opinion/blogs/pat-garofalo/2013/11/25/the-importance-of-switzerlands-112-ceo-pay-cap-vote

Thompson, P., & Pronsky, J. (1975). Secrecy or disclosure in management compensation?. *Business Horizons*, 18(3), 67-74.

Varma, A., & Stroh, L. K. (2001). The impact of same-sex LMX dyads on performance evaluations. *Human Resource Management*, 40(4), 309-320.

Vroom, V. (1964). Work and motivation. New York, NY: John Wiley.

Walster, E., Berscheid, E., & Walster, G. W. (1973). New directions in equity research. *Journal of Personality and Social Psychology*, 25(2), 151.

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems*, 12(4), 5-33.

Wayne, S. J., & Liden, R. C. (1995). Effects of impression management on performance ratings: A longitudinal study. *Academy of Management journal*, 38(1), 232-260.

Williamson, O. E. (1993). Calculativeness, trust, and economic organization. *The Journal of Law and Economics*, 36(1, Part 2), 453-486.

Wong-On-Wing, B., Guo, L., Li, W., & Yang, D. (2007). Reducing conflict in balanced scorecard evaluations. *Accounting, Organizations and Society*, 32(4), 363-377.

Wood, R. E. (1986). Task complexity: Definition of the construct. *Organizational Behavior and Human Decision Processes*, 37(1), 60-82.

Woods, A. (2012). Subjective adjustments to objective performance measures: The influence of prior performance. *Accounting, Organizations and Society*, 37(6), 403-425.

WorldatWork & Vivient Consulting. (2012). Private Company Incentive Pay Practices Survey. http://www.worldatwork.org/waw/adimLink?id=58146

Yim, A. T. (2001). Renegotiation and relative performance evaluation: Why an informative signal may be useless. *Review of Accounting Studies*, 6(1), 77-108.

Zijlstra, F. R., Roe, R. A., Leonora, A. B., & Krediet, I. (1999). Temporal factors in mental work: Effects of interrupted activities. *Journal of Occupational and Organizational Psychology*, 72(2), 163-185.

Doctoral dissertations from the Faculty of Economics and Business, see:
http://www.kuleuven.ac.be/doctoraatsverdediging/archief.htm.